

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 25, 2003, 04:01:26 ; Search time 4877 Seconds  
(without alignments)  
4105.542 Million cell updates/sec

Title: US-09-931-007A-1  
Perfect score: 3508  
Sequence: 1 MGETLDSPIDPSDSFTDR.....KTEEDMSLHPLDIEYKDLX 688

Scoring table:  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2054640 seqs, 14551402878 residues  
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DEV=xjh  
-O=/cgr2\_1/USPTO.spool/US09931007/runat\_14022003\_102553\_4678/app.query.fasta\_1.839  
-DB=GenDb1 -QPM=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=-1 -MATRIX=Blasum62 -TRANS=human40.cdi -LIST=45  
-DOCALLION=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFM=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931007 @CGN 1.1.3052.ernat.14022003.102553.4678 -NCPu=6 -ICPU=3  
-NO\_XIPXY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV\_TIMEOUT=120  
-MARK\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl : \*  
1: gb\_da : \*  
2: gb\_hgt : \*  
3: gb\_in : \*  
4: gb\_cm : \*  
5: gb\_ov : \*  
6: gb\_pat : \*  
7: gb\_ph : \*  
8: gb\_pl : \*  
9: gb\_pr : \*  
10: gb\_ro : \*  
11: gb\_sts : \*  
12: gb\_sy : \*  
13: gb\_un : \*  
14: gb\_vi : \*  
15: em\_da : \*  
16: em\_fun : \*  
17: em\_hum : \*  
18: em\_in : \*  
19: em\_mu : \*  
20: em\_om : \*  
21: em\_or : \*  
22: em\_ov : \*  
23: em\_pat : \*  
24: em\_ph : \*  
25: em\_pl : \*  
26: em\_ro : \*  
27: em\_sts : \*  
28: em\_un : \*

29: em\_vi : \*  
30: em\_htg\_hum : \*  
31: em\_htg\_inv : \*  
32: em\_htg\_other : \*  
33: em\_htg\_mus : \*  
34: em\_htg\_pln : \*  
35: em\_htg\_rtd : \*  
36: em\_htg\_man : \*  
37: em\_htg\_vtl : \*  
38: em\_sy : \*  
39: em\_htgo\_hum : \*  
40: em\_htgo\_mus : \*  
41: em\_htgo\_other : \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2231	63.6	1518	9 HS063415	U63415 Human perox
2	2231	63.6	1608	6 AR121467	AR121467 Sequence
3	2231	63.6	1608	9 HS079012	U79012 Human ligan
4	2217	63.2	1759	9 AF033103	AF033103 Macaca mu
5	2213	63.1	1765	9 AY048695	AY048695 Macaca fa
6	2196.5	62.6	1679	6 AR203332	AR203332 Sequence
7	2196.5	62.6	1679	9 HUMPEPARG	DB3233 Homo sapien
8	2186	62.3	1849	9 AY048698	AY048698 Macaca fa
9	2172.5	61.9	1539	4 AF059245	AF059245 Sus scrofa
10	2171.5	61.9	1758	4 AF103946	AF103946 Sus scrofa
11	2160.5	61.6	1640	4 SSC6757	AJ066757 Sus scrofa
12	2156	61.5	1782	10 BC021798	BC021798 Mus muscu
13	2153	61.4	1780	10 MM009138	U09138 Mus musculu
14	2148	61.2	1518	10 AB019561	AB019561 Rattus no
15	2148	61.2	1785	10 AF156666	AF156666 Rattus no
16	2143	61.1	1570	10 RNPPARGM2	Y12882 Rattus norv
17	2136.5	60.9	2028	9 AY048697	AY048697 Macaca fa
18	2132	60.8	1815	4 BTPPARG2	Y12420 B. taurus nr
19	2100	59.9	1766	9 HSPPARGM	X90563 H. sapiens m
20	2100	59.9	1808	9 HUMPEPARG	LC0904 Homo sapien
21	2095.5	59.7	1767	9 BC006811	BC006811 Homo sapi
22	2080	59.3	1766	9 AF033343	AF033343 Macaca mu
23	2080	59.3	1794	9 AF033342	AF033342 Macaca mu
24	2076	59.2	1800	9 AY048694	AY048694 Macaca fa
25	2076	59.2	2063	9 AY048699	AY048699 Macaca fa
26	2075	59.2	1713	9 AY048696	AY048696 Macaca fa
27	2074	59.1	1884	9 AY048700	AY048700 Macaca fa
28	2065.5	58.9	1811	6 AX409565	AX409565 Sequence
29	2065.5	58.9	1844	6 AR139020	AR139020 Sequence
30	2065	58.9	2005	6 AR030509	AR030509 Sequence
31	2062	58.8	1636	10 AB011365	AB011365 Rattus no
32	2062	58.8	1798	10 AF246457	AF246457 Rattus no
33	2062	58.8	1838	10 AF246458	AF246458 Rattus no
34	2062	58.8	1848	10 AF156665	AF156665 Rattus no
35	2062	58.8	2130	10 MM010374	U10374 Mus musculu
36	2054	58.6	1428	10 MM010374	U10374 Mus musculu
37	2054	58.6	1709	10 CGPPARGRN	Z30972 C. griseus m
38	2053.5	58.5	1630	4 SSC6756	AJ066756 Sus scrofa
39	2047	58.4	1759	4 OC084893	U84893 Oryctolagus
40	2045	58.3	1428	10 AF317514	AF317514 Cavia por
41	2045	58.3	1711	4 BTPPARG1	Y12419 B. taurus nr
42	2035	58.0	1800	10 U01664	U01664 Mus musculu
43	2028	57.8	1428	5 AF464684	AF464684 Anas plat
44	1933	55.1	2089	5 AB045597	AB045597 Gallus ga
45	1933	55.1	2089	5 AF163811	AF163811 Gallus ga

RESULT 1

ALIGNMENTS

HS063415 1518 bp mRNA linear PRI 18-JUL-1996  
LOCUS HS063415  
DEFINITION Human peroxisome proliferator activated receptor gamma 2 mRNA,  
complete cds.  
ACCESSION U63415  
VERSION U63415.1 GI:1432176  
KEYWORDS  
SOURCE Homo sapiens.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 1518)  
Elbrecht, A., Chen, Y., Cullinan, C.A., Hayes, N., Leibowitz, M.D.,  
Moller, D.E. and Berger, J.  
TITLE Molecular cloning, expression and characterization of human  
peroxisome proliferator activated receptors gamma 1 and gamma 2  
Biochem. Biophys. Res. Commun. (1996) In press  
JOURNAL 2 (bases 1 to 1518)  
REFERENCE Elbrecht, A., Chen, Y., Cullinan, C.A., Hayes, N., Leibowitz, M.D.,  
Moller, D.E. and Berger, J.  
AUTHORS Direct Submission  
TITLE Submitted (10-JUL-1996) Molecular Endocrinology, Merck Research  
JOURNAL Labs., P.O. Box 2000, Mail Drop R807-265, Rahway, NJ 07060, USA  
FEATURES  
source Location/Qualifiers  
1..1518  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/cell\_type="fat cell"  
1..1518  
/codon\_start=1  
/product="peroxisome proliferator activated receptor gamma  
2"  
/protein\_id="AAB04028.1"  
/db\_xref="GI:1432177"  
/translation="MGETLGDSPTDPDSSTDTLNLANSCEMTWDTMPFPPNNG  
ISVYDLVMEHSHSFDIKPTTVDFFSISITPHIEDIPEPTDENVADYKDLQEX  
QSAIKVEPASPSPYSEKTOLNKPHEBSPNSLHIECVCDKASGFHYGVHACEGK  
GFFRRITRLKLIYRCDLNCRIHKRSKRCQYCFQCLAVGSHNAILRFRMQAEK  
EKLAETISSDIDOLNPESADRLAKHLHYDSIKSPITTKAAAILGKTTDSPEY  
IYDMSLSMGEDKIKFKHITPILOESKEVARIRCOGQREVEAORTETAKSIPG  
VNDLNDQVTLTKYGVHEIITMTLASLNKGVILSEGGFTMEFLKSKPPGDEM  
EPKFEFAKFNALDSDSLATFIAVILISDGRGLNKFIEDIIDLQALDLQK  
LNHPSSQLFKLLQKMTDLNROIYTEHVQLLVTKTETDMSLHPLOEIKDY"

BASE COUNT 438 a 363 c 346 g 371 t  
ORIGIN

Alignment Scores:  
Pred. No.: 1.32e-159 Length: 1518  
Score: 2231.00 Matches: 443  
Percent Similarity: 87.72% Conservative: 0  
Best Local Similarity: 87.72% Mismatches: 0  
Query Match: 63.60% Indels: 62  
DB: 9 Gaps: 1

US-09-931-007a-1 (1-688) x HS063415 (1-1518)

OY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20  
|||||  
DB 1 ATGGGTGAACCTCTGGAGATTCTCTTATGACCCAGAAAGGATTCCTTCACTATACA 60  
|||||

OY 21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGluMetProPheThrPro 40  
|||||  
DB 61 CTGCTGCAACATATCAACAAGATGACCATGTTGACACAGAGATGCATCTGGGCC 120  
|||||

OY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
|||||  
DB 121 ACCAAGCTTGGATCAGCTCCGATCTCTCCGTAATGGAAGACCATCCACCTCTTT 180  
|||||

OY 61 AspIleIleProPheThrThrValAspPheSerIleSerThrProHisIleGluAsp 80  
|||||  
DB 181 GATATCAAGCCCTTCTACTACTGTTGCTCCACATTTCTACTCCACATTACGAGAAC 240  
|||||

OY 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGluIn 100  
|||||

|||||  
DB 241 ATTCATTACAAAGAACATCCAGTGGTTGCAGATTCAAGATGATGACCTGAACCTCA 300  
|||||

OY 101 GUTYrGlnSerAlaIleIleYsValGluProAlaSerProProTyrTyrSerGluYsThr 120  
|||||

DB 301 GAGTACCAAGAGCAATCAATCAAGTGGAGCCTGCATCTCCACCTTATATTCAGAAAGACT 360  
|||||

OY 121 GlnLeuTyrAsn----- 124  
|||||

DB 361 CAGCTCTACAAATACCTTCATATGAAGCCTTCACCTCCATGCGCAATTGATGTGCT 420  
|||||

OY 124 ----- 124

DB 421 GTCTGTGAGATAAAGCTTCGATTTTCATGTGAGATTGATGCTTGGAAGATGCAAG 480  
|||||

OY 124 ----- 124

DB 481 GGTTCCTCCGAGAACATCATGATTGAAGCTTATCTATGACAGATGATCTTAACGT 540  
|||||

OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
|||||

DB 541 CGGATCCACAAAAAGTAGAATATTAATGTCAGTCTGCTGCTTCAGAAATGCCCTTCA 600  
|||||

OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluYsGluYs 158  
|||||

DB 601 GTGGGAGATGTCATTAATGCCATCAGGTTTGGCGGATGCCAGCGCAGAGAGAGAG 660  
|||||

OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
|||||

DB 661 CTGTTGGGAGATCTCCAGTATATCCAGCCAGCTGAATCCAGATCCGCTGCCTCCGG 720  
|||||

OY 179 AlaLeuAlaYsHisLeuTyrAspSerTyrIleYsSerPheProLeuThrLysAlaYs 198  
|||||

DB 721 GCCCTGGCAAAACATTTGTATGACTCATACATTAAGTCTCTTCCCTACCAAGCAAG 780  
|||||

OY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
|||||

DB 781 GCGAGGCGGATCTTGACAGGAAGAACACAGCAATCCATTCGTTATCATGATCATG 840  
|||||

OY 219 AsnSerLeuMetMetGlyGluAspLysIleYsPheYsHisIleThrProLeuGlnIle 238  
|||||

DB 841 AATTCCTTAATGATGGAGAAATATAATCAAGTTCAACATCACACCCTCGAGAG 900  
|||||

OY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258  
|||||

DB 901 CAGACCAAGAGGTGGCATCCATCCGATCTTTCAGGGCTGCCAGATTCCGCTCGAGAGCT 960  
|||||

OY 259 ValGlnIleIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
|||||

DB 961 GTGCAGAGATCACAGAGTATGCCAAAGCATTCCTGGTTTGTAAATCTGACTTGAAAC 1020  
|||||

OY 279 AspGlnValIleThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
|||||

DB 1021 GACCAAGTAACCTCTCCAAATATGAGATCCACAGATCATTTACAAATCTGGCCTCC 1080  
|||||

OY 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318  
|||||

DB 1081 TTGATGATTAAGATGGGGTCTTCATATCCGAGGCCAAAGCTTCATATCAAGGAGATT 1140  
|||||

OY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338  
|||||

DB 1141 CTAAAGACCTGCGCAAGCCTTTGTGACTTTATGAGGCCCAAGTTGAGATTCTCTG 1200  
|||||

OY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
|||||

DB 1201 AAGTTCATATGACATGGAATATGATACAGCGACTTGCGCAATATTAATGCTGCTTAATT 1260  
|||||

OY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
|||||

DB 1261 CTCAGTGGAGACCGCCAGGTTTGTGTAATGTAGAGCCCATTTGAAGACATTCAGACACAC 1320  
|||||

OY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398  
|||||

Db 1321 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGAGTCTCAGACGCTGTT 1380  
Qy 399 AAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIThrGlnHisValGlnLeu 418  
Db 1381 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACGATTGTGACGAGACACGCTGCGACTA 1440  
Qy 419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIle 438  
Db 1441 CTGCAGGTGATCACAAGAACGAGACAGACATGAGTCTTCACCCGCTCCTGCAGAGATC 1500  
Qy 439 TyrLysAspLeuTyr 443  
Db 1501 TACAAAGACTGTAC 1515

RESULT 2  
AR121467

LOCUS AR121467 1608 bp DNA linear PAT 16-MAY-2001  
DEFINITION Sequence 3 from patent US 6159734.  
ACCESSION AR121467  
VERSION AR121467.1 GI:14105043

KEYWORDS  
SOURCE .  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1608)  
AUTHORS McKay, R., Borchers, A. H. and Baker, B. F.  
TITLE Antisense modulation of peroxisome proliferator-activated receptor  
gamma expression  
JOURNAL Patent: US 6159734-A 3 12-DEC-2000;  
FEATURES Location/Qualifiers  
source 1..1608  
/organism="unknown"

BASE COUNT 462 a 380 c 361 g 405 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 1,416-159 Length: 1608  
Score: 2231.00 Matches: 443  
Percent Similarity: 87.72% Conservative: 0  
Best Local Similarity: 87.72% Mismatches: 0  
Query Match: 63.60% Indels: 62  
DB: Gaps: 1

US-09-931-007A-1 (1-688) x AR121467 (1-1608)  
Qy 1 MetGlyGlnThrLeuGlnLysSerProIleAspProGluSerAspSerPheThrAspThr 20  
Db 91 ATGGGTAAGACTCTGGAGATTCTCTATTGACCCAGAAAGGATTCCTTCACTGATACA 150  
Qy 21 LeuSerAlaAsnIleSerGlnLumethrMetValAspThrGluMetProPheThrPro 40  
Db 151 CTGCTGCAACATATCACAAGAAATGACCATGCTTGACACAGACATGCTTGGCCC 210  
Qy 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
Db 211 ACCAAGCTTGGGATGAGCTCCGCTGATCTCTCGTAATGAGAACACCTCCACTCTCTT 270  
Qy 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80  
Db 271 GATATCAACCCCTCAGTACTGTGATCTTCACACATTCTACATTCACATTAAGAGAC 330  
Qy 81 IleProPheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
Db 331 ATTCATTACACAAGAACGATCCAGTGGTTCAGATTAACAATAGACTGAACCTTCA 390  
Qy 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
Db 391 GAGTACCAAGAGCAATCAAAAGTGGAGCTGCATCTCCACCTTATTATCTGAGAAAGACT 450  
Qy 121 GlnLeuTyrAsn 124  
Db 451 CAGCTCTACAATTAAGCTCATGAAGAGCCTTCAACTCCCTCATGAGCAATTGAATGTCT 510

Qy 124 ----- 124  
Db 511 GTCTGTGAGATAAAGCTTTCGATTTCATATGAGAGTTCATGCTTGAAGATGCAAG 570  
Qy 124 ----- 124  
Db 571 GGTTCCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGATCTTAACGT 630  
Qy 125 ----- ArgAsnLysCysGlnTyrCysAspPheGlnLysCysLeuIle 138  
Db 631 CGGATCCACAAAAAAGTGAATTAATGATCAGTACTGCGGTTTCAGAAATGCCCTTCA 690  
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158  
Db 691 GTGGGATGTCTCATTAATGCCATCAGTTTGGCGGATGCCACAGCCGAGAGAGAGAG 750  
Qy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
Db 751 CTGTGGCGGAGATCTCCAGTATGATCCAGCAGCTGAATCCAGAGTCCGCTGACCTCCG 810  
Qy 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198  
Db 811 GCCCTGGCAAAACATTTGATGACTCATATCAATAGTCCCTCCAGCAAAAGCAAG 870  
Qy 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
Db 871 GCGAGGCGGATCTTACAGAGAAAGACACAAATCACCATTGCTATCATGACATG 930  
Qy 219 AsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238  
Db 931 AATTCCTTAATGATGGAGAGATTAATCAAGTTCAAAACATCACCCTCCGAGAG 990  
Qy 239 GlnSerLysGluValAlaIleArgIlePheGlnGlnCysGlnPheArgSerValGluAla 258  
Db 991 CAGAGCAAGAGAGTGGCCATCCGATCTTTCAGGGCTCCAGTTCCGCTGGAGAGCT 1050  
Qy 259 ValGlnGluIleThrGlyTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
Db 1051 GTGCAGAGATCAGACAGATATGCCAAAAGCATCTCGTTTGTAACTTGACTTGAAAC 1110  
Qy 279 AspGlnValIleThrLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
Db 1111 GACCAAGTAAGCTCTCAAAATATGAGTCCACAGATCATTTACAAATCTGCTGCC 1170  
Qy 299 LeuMetAsnLysAspGlyValLeuIleSerGluGlnGlnGlyPheMetThrArgGluPhe 318  
Db 1171 TTGATGAATTAAGATGGGTTCTCTATCCAGGGCCAAAGGCTTCATGACAAGGAGCTTT 1230  
Qy 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338  
Db 1231 CTAAAGACCTCCGAAAGCCTTTTGTGATGATGAGCCCAATTTGAGAGTTGAGTTGCTGTG 1290  
Qy 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
Db 1291 AAGTTCAATGACACTGGAATTTGATGACAGCGACTGGCAAAATTTATGCTGTCTATT 1350  
Qy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
Db 1351 CTCAGTGGAGACCGCCGAGTTTGTGAATGTGAAGCCCATTTGAAGACATTCACAAGACAC 1410  
Qy 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLysAsnHisProGluSerSerGlnLeuPhe 398  
Db 1411 CTGCTACAAAGCCCTGGAGCTCAGCTGAAGTGAACCAACCTGAGTCTCAGAGTGT 1470  
Qy 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIThrGlnHisValGlnLeu 418  
Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTCAGAGAACAGCTGAGACTA 1530  
Qy 419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIle 438  
Db 1531 CTGCAGGTGATCACAAGAACGAGACAGATGATGCTTTCACCCGCTCCTGCAGAGATC 1590  
Qy 439 TyrLysAspLeuTyr 443





Db 1291 AAGTTCATGACCTGATAGATGACAGCAGCTTGCCAAATATTATTGCTCATTAATT 1350

Qy 359 LeuSerGIyAspArgProGIyLeuLeuAsnValIySProlleGIuAspIleGIuAspAsn 378  
|||||

Db 1351 CTCACATGGAGACCGCCAGGTTTGTGATGTGAAGCCCATTTGAAGACATTTAAACAAAC 1410  
|||||

Qy 379 LeuLeuGIuAlaLeuGIuLeuGIuLeuGIuLeuAsnHisProGIuSerSerGIuLeuPhe 398  
|||||

Db 1411 CTGCTACAAAGCCCTGAGCTCCAGCTGAAGCAACCCCTGAGTCTCCACAGCTGTTT 1470  
|||||

Qy 399 AlaIySLeuLeuGIuLeuMethThrAspLeuArgGIuIleValThGIuHisValGIuLeu 418  
|||||

Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTCACGGAACACCTGAGCTTA 1530  
|||||

Qy 419 LeuGIuAlaIleIySLeuThrcGIuThrAspMetSerLeuHisProLeuLeuGIuIle 438  
|||||

Db 1531 CTGCAAGGTGATTAAGAAAGAGAGACAGACATGAGTCTTACCCCTCTCGAGAGATC 1590  
|||||

Qy 439 TyrIyAspLeuTyr 443  
|||||

Db 1591 TACAAGAGACTGTAC 1605

RESULT 4

AF033103 1759 bp mRNA linear PRI 27-NOV-1997

LOCUS Macaca mulatta peroxisome proliferator-activated receptor gamma 2

DEFINITION (PPARG2) mRNA, complete cds.

ACCESSION AF033103

VERSION AF033103.1 GI:2645730

KEYWORDS

SOURCE Macaca mulatta

ORGANISM Macaca mulatta

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Macaca.

REFERENCE 1 (bases 1 to 1759)

AUTHORS Hotta,K., Gustafson,T.A., Yoshioka,S., Ortmeier,H.K., Bodkin,N.L. and Hansen,B.C.

TITLE Relationships of PPARG and PPARG2 mRNA levels to obesity, diabetes, and hyperinsulinemia in rhesus monkeys

Unpublished

2 (bases 1 to 1759)

Hotta,K., Gustafson,T.A., Yoshioka,S., Ortmeier,H.K., Bodkin,N.L. and Hansen,B.C.

Direct Submission

Submitted (05-NOV-1997) Department of Physiology, University of Maryland, School of Medicine, 10 South Pine Street, MSTF Room 6-00, Baltimore, MD 21201, USA

FEATURES

source Location/Qualifiers

1..1759

/organism="Macaca mulatta"

/db\_xref="taxon:9544"

/tissue\_type="adipose"

1..1759

/gene="PPARG2"

25..1542

/gene="PPARG2"

/note="transcription factor"

/codon\_start=1

/product="peroxisome proliferator-activated receptor gamma 2"

/protein\_id="AAB87480.1"

/db\_xref="GI:2645731"

/translation="MGETIGDSPIDPESDFTDLSANISOETIMVDTMPMPPTNG ISSYDLSVMDHSHSPDIKPTPTVDPSSISAPHEDI.PFTPTDMVADKYDLKLOX OSAIKVEPASPPIYSEKTOLYNKPHEPNSSLMIEKRVGDKASGHVGNACGCK EKLAEISSDQNPESADLRALAKHLUDYSIKSFPLTAKARAILGTGTTDKSPV IYDNLNLMGDEKIKFKHITPLOSKEVAIRIFOGCFRSEVAOETEAQKIPF VNLIDNVLTGKYGHETIYTMASLMNKGVLISGCGFMTEFLSKRPKGDMD EPRFEPAVKFNAILDSDLAIFAVIILSDRGLNVRPIEDIONLQALFLQK LNHPESSOLFKLQKMTDLKQIVTEHVQLQVTKTETDMSLHPLOETIKDLY"

BASE COUNT 532 a 407 c 376 g 444 t

ORIGIN

Alignment Scores:

Pred. No.: 1..8e-158 Length: 1759

Score: 2217.00 Matches: 440

Percent Similarity: 87.08% Conservative: 5

Best Local Similarity: 86.11% Mismatches: 4

Query Match: 63.20% Indels: 62

DB: 9 Gaps: 1

US-09-931-007a-1 (1-688) x AF033103 (1-1759)

Qy 1 MetGIyGurThLeuGIyAspSerProIleAspProGIuSerAspSerPheThAspThr 20  
|||||

Db 25 ATGGGTAAACTGTGGGAGATTCCTCTATGACCAGAAAGCATTCCTCAGTATGACA 84  
|||||

Qy 21 LeuSerAlaAsnIleSerGIuMetThrcValAspThrGIuMetProPheThrPro 40  
|||||

Db 85 CTGCTCGCAACATATTCACAGAAATTAACATGTTGACACAGAGATGCCATTCTGGCCC 144  
|||||

Qy 41 ThrAspMetGIyIleSerSerValAspLeuSerValMetGIuAspHisSerHisSerPhe 60  
|||||

Db 145 ACCAATTGGGATCGCTCCGTGATCTCTCCGTAATGACGACCACTCCACTCTTT 204  
|||||

Qy 61 AspIleIySProlleThrThValAspPheSerSerIleSerThrProHisIleTyrGIuAsp 80  
|||||

Db 205 GACATCAAGCCCTTCACAGCTGTTGACTTCTCCAGCATTCTGCTCCCATTTAGAACAC 264  
|||||

Qy 81 IleProPheThrArgThrAspProValAlaAspTyrIyTyrAspLeuIySLeuGIu 100  
|||||

Db 265 ATTCATTCACAGAACAGATCCGATGCTCGGATTAATGATGACCTGAACCTTCAA 324  
|||||

Qy 101 GIuTyrGIuSerAlaIleIySValGIuProIleAspProIyTyrSerGIuIySThr 120  
|||||

Db 325 GAGTACCAAGAGTCATCAATGAAGGAGCTGCTTCCACCTTATTTATTCGAGAGACT 384  
|||||

Qy 121 GIuLeuTyrAsn 124  
|||||

Db 385 CAGCTCTCAATTAAGCTCAGAGAGAGCTTCCACCTCCATGGAATGATGCTGT 444  
|||||

Qy 124 124  
-----

Db 445 GTCTCGGAGATAAGCTTGTGATTCACATGAGATTCACTGTGGAAGATGCAAG 504  
-----

Qy 124 124  
-----

Db 505 GCTTTCTCCGAGACATCAGATTGAGCTTATCTATGACAGATGATGATCTTAAGTCT 564  
-----

Qy 125 -----ArgAsnIyScysGIuTyrCysArgPheGIuIyScysLeuAla 138  
|||||

Db 565 CGGATCCACAAAAGAAATTAATGTCAGTACTGTGCGTTTCAAAATGCTTGA 624  
|||||

Qy 139 ValGIuMetSerHisAsnAlaIleArgPheGIuArgMetProGIuAlaGIuIySLeuIyS 158  
|||||

Db 625 GTGGGAGATGCTCAATTAAGCCATCAGGTTTGGGAGCCACAAAGCCGAGAGAGAG 684  
|||||

Qy 159 LeuLeuAlaGIuIleSerSerAspIleAspGIuLeuAsnProGIuSerAlaAspLeuArg 178  
|||||

Db 685 CTGTTGGCGGAGATCTCCAGTATATGACACAGCTGAGCCAGAGATCGCTGACCTCCGG 744  
|||||

Qy 179 AlaLeuAlaIySHisLeuTyrAspSerTyrIleIySLeuPheProLeuThrIySAlaIyS 198  
|||||

Db 745 GCCCTGGCAAAACATTTGTATGACATCATACATAAAGTCCCTCCGCTGACCAAGCAAG 804  
|||||

Qy 199 AlaArgAlaIleLeuThrcGIyLysThrThrAspIySLeuSerProPheValIleTyrAspMet 218  
|||||

Db 805 GCGAGGGGAGATCTTGACAGAGAAAGACAAACAATACCATTCGTTATATGACATG 864  
|||||

Qy 219 AsnSerLeuMetMetGIuAspIySLeuIySLeuPheIySHisIleThrProLeuGIuIyS 238  
|||||

Db 865 AATTCCTTAATGATGGAGAAATTAATCAAGTTCAACACATCACCCCACTGAGAGAG 924  
|||||

Qy 239 GIuSerIySLeuValAlaIleArgIlePheGIuIySLeuIySLeuPheArgSerValGIuAla 258  
|||||

|||||  
Db 925 CAGACGAAGAGGTGGCCATCCGACATCTTTCAGAGGCTCGCATTCGCTCCGTGGAGGCT 984  
OY 259 VALGNGLIUlllthrgluThrAlaLysSerlleProGlyPheValAsnLeuAspLeuAsn 278  
Db 985 GGGCAGGAGATCACAGAGATGCGCAAAAGCATTCCTGGTTTGGTAAATCTTGACATTTGAAC 1044  
OY 279 AAGGlnValThrLeuLeuLysTyrGlyValHisGluIlelleTyrThrMetLeuAlaSer 298  
Db 1045 GACCAAGTAACTCTCTCCAATATATGAGTCATGATCATTTTACACATTCGCTGGCTCC 1104  
OY 299 LeuMetAsnLysAspGlyValLeuIleSerGlyGlyGlyPheMetThrArgGluPhe 318  
Db 1105 TTGATGAAATAAAGATGGGCTTCTCATATCCGAGGGCCAGAGCTTCATGACAAGGGAATT 1164  
OY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338  
Db 1165 CTAAAGAGCTCGCAAGGCTTTGGTGTGACTTATGAGACCCCAAGTTTGAATTTGCTGTG 1224  
OY 339 LysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
Db 1225 AAGTTCAATGCATCGAATATGATGACAGCATTCGGCAATTTATTCCTGTCATTAAT 1284  
OY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
Db 1285 CTCAGTGGAGATCGCCAGGTTTGTGATGTGAAGCCCATTTGAAGACATTCACAAGCAAC 1344  
OY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGlnLysSerGlnLeuPhe 398  
Db 1345 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGTGAGTCTCGAGCTGTTT 1404  
OY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418  
Db 1405 GCCAAGCTCTCTCAGAAATGACAGACCTCAGACAGATTGTCACGGAACACGTGCAGCTA 1464  
OY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIle 438  
Db 1465 CTGCGAGTGATCAGAGAGCGAGAGACATGAGCTTCACCCGCTCTCGCAGAGGATC 1524  
OY 439 TTYLysAspLeuTyrAlaThrPalaIleLeuThr 449  
Db 1525 TACAAAGACTTGTACTAGCAGAGAGTCTGTGAC 1557

RESULT 5  
LOCUS AY048695 1765 bp mRNA linear PRI 21-MAY-2002  
DEFINITION Macaca fascicularis peroxisome proliferator-activated receptor  
gamma 2 (PPARGgamma2) mRNA, complete cds.  
ACCESSION AY048695  
VERSION AY048695.1 GI:15723731  
KEYWORDS  
SOURCE  
ORGANISM Macaca fascicularis.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
Cercopithecidae; Macaca.  
REFERENCE  
AUTHORS Zhou, J., Wilson, K.M. and Medh, J.D.  
TITLE Genetic analysis of four novel peroxisome proliferator activated  
receptor gamma splice variants in monkey macrophages  
JOURNAL Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)  
MEDLINE 12054596  
PUBMED 22050014  
JOURNAL 22050014  
TITLE Direct Submission  
AUTHORS Zhou, J., Wilson, K.M. and Medh, J.D.  
JOURNAL Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,  
200 Hawkins Drive, Iowa, IA 52242, USA  
FEATURES  
source  
1..1765  
/organism="Macaca fascicularis"  
/db\_xref="taxon:9541"  
/cell\_type="macrophage"

gene 1..1765  
/gene="PPARGgamma2"  
25..1542  
CDS  
/gene="PPARGgamma2"  
/note="transcription factor"  
/codon\_start=1  
/product="peroxisome proliferator-activated receptor gamma  
2"  
/protein\_id="AA05262.1"  
/db\_xref="GI:15723732"  
/translation="MGETLGDSPIDPESDSTDTLSANISQETWVTEMPFPTNG  
/ISVDLSVMDHSHFDIKPPTVDFESSISAPHEEDIPETFDPMVADYKDKLQLEY  
QSALKEVBPSPYSEKTLQYKPHPEPSNLSIAEIVCGDKASGFHGVHACGCK  
GFPRTRIRKLIVRCDLNCRIHKSRNKCQYCRFOCLAVGSHNAIRGMRQAEK  
EKLAEISSDIDOLNPESADLRALAKHLUDSYIKSPFLTKAKARAILTGKTDKSPV  
IYDMSLSMIGBDKIKPKHPIPLQROSKVEARIRGCGQFSEVAVQETVAKSIRPF  
VNLNDQVTLKATGVHEIITMLASLMNKDGVITSGQGFMTREFLSKLRPGDPM  
EPKFEFAVFNALLEDSDLAIFLAVIILISDRGLNLVPIEDIQNLQALQELQK  
LNHPSSQLFVKLLQKMTDLROIYTHEVOLLDQVIKKTETDMSLHPLQEIYKDLX"  
BASE COUNT 537 a 409 c 376 g 443 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,61e-158 Length: 1765  
Score: 2213.00 Matches: 439  
Percent Similarity: 86.89% Conservative: 5  
Best Local Similarity: 85.91% Mismatches: 5  
Query Match: 63.08% Indels: 62  
D: 9 Gaps: 1  
US-09-931-007a-1 (1-688) x AY048695 (1-1765)  
OY 1 MetGlyGluThrLeuGluYAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
Db 25 ATGGGTGAACACTCTGGAGATTCCTCTATGTGCCAGAAAGGATTCCTTCACTATACA 84  
OY 21 LeuSerAlaAsnIleSerGlnLumethrMetValAspThrGluMetProPheThrPro 40  
Db 85 CTGCTGCAAAACATATCACAAAGAAATTCAGATGTTGACACAGAGATGCCATTTGGCCC 144  
OY 41 ThrAsnPhleglyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
Db 145 ACCAAGCTTTGGGATGAGCTCCGCGATCTCTCGTAATGAGAGACCACTCCACCTCTTT 204  
OY 61 AspIleLysProPheThrThrValAspPheSerIleSerThrProHisTyrGlnAsp 80  
Db 205 GACATCAACCCCTTCACTACTGCTTACTTCTCACACATTTCTGCTCCACATTAAGAAC 264  
OY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
Db 265 ATTCATTCACAAAGAACAGATCCGATGGTGGGATTTAAGTAGACCTGAACTTCAA 324  
OY 101 GluTyrGlnSerAlaIleLysValIleProAlaSerProProTyrTyrSerGluLysThr 120  
Db 325 GAGTACCAAGAGCAATCAAAAGTGAGCGCTGCTTCCACCTATTATTATTCGAGAAACT 384  
OY 121 GlnLeuTyrAsn 124  
Db 385 CAGCTCTACATTAAGCTTCACAGAGAGCTTCCAACTCCCTCATGCGCAATGCAATGTCGT 444  
OY 124 124  
Db 445 GTCTCGGAGATAAGCTTCTGGATTCACATATGAGATTCATGCTTGGAAGATGCAAG 504  
OY 124 124  
Db 505 GGTTCCTTCGAGAGAACATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAAGTGT 564  
OY 125 -----ArgAsnLysCysGlnTyrCysAspGlnLysCysLeuAla 138  
Db 565 CGGATCCACAAAAAAGTAGAATTAATGTCACTGTGCTTTCACAAATGCTTCA 624  
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158

|||||  
625 GTGGGATGTCATATATGATCCATGAGTTGGGCGGATGCCAAGCCGAGAGAGAG 684  
159 LeuLeuAlaIleLeuThrSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
665 CTGTGGGGGAGATCTCCAGATATCGACAGCGAAGCCAGATCCGCTGACCTCCGG 744  
179 .AlaLeuAlaIleLeuThrSerAspSerIleLeuSerPheProLeuThrLysAlaLys 198  
745 GCCCTGGGAAAAATTTGTATGACATACATAAAGTCTTCCCTCCGCTGACCAAGAG 804  
199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
805 GCGAGGCGGATCTTACAGAGAAAGACAGACAAATACCATTCGTTATCTATGACATG 864  
219 AsnSerLeuMetMetGlyGlnAspLysIleLeuSerPheLysHisIleThrProLeuGlnIle 238  
865 AATTCTTAATGATGAGAGATAAATCAAGTTCAAAACATCAACCCCTGACAGAG 924  
239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258  
925 CAGACCAAGAGAGTGTATCCGCAATCTTTCAGGGCTCCAGATTGCGTCCGTGGAGCT 984  
259 ValGlnIleIlePheGlnTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
985 GTGCGAGAGATCACAGAGATATGCCAAAGCATTCCTGTTCTGTAAATCTTGACTTGAC 1044  
279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
1045 GACCAAGTACTCTCTCAATATGAGATGCATGAGATCATTTACATATCTGCCCCC 1104  
299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318  
1105 TTGATGATTAAGATGAGGGGTTTCTCATATCCAGAGCCCAAGCTTCATACAGGAATTT 1164  
319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338  
1165 CTAAAGAGCCCTGCGAAAGCCTTTTGTGACTTTATGAGCCCAAGTTGAGTTGCTGTG 1224  
339 LysPheAsnAlaLeuGlnLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIle 358  
1225 AAGTTCAATGCACTGTGAAATGTGATACAGCATTTGGCAAAATTTATTTGCTGCTATTATT 1284  
359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
1285 CTCATGTGAGATCGCCGAGGTTTGTGTAATGTGAAGCCCATTTGAAGACATTCAGACACAC 1344  
379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398  
1345 CTGCTACAGAGCCCTGAGCTCCAGCTCAAGCTGAACCACTGAGTCTCGCAGCTGTTT 1404  
399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGlnHisValGlnLeu 418  
1405 GCCAAGCTGCTCCGAAATAATACAGACCTTCAGACAGATTGTACAGGAACAGCTGACAGTA 1464  
419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnIleIle 438  
1465 CTGCGAGTGATCAGAAACAGAGACGACATGACTTTCACCCGCTCTGACAGAGATC 1524  
439 TyrLysAspLeuTyrAlaIleThrAlaIleLeuThr 449  
Db 1525 TACAAGGACTGTACTACAGAGAGTCTGAGC 1557

RESULT 6  
AR203332 1679 bp DNA linear PAT 20-JUN-2002  
LOCUS AR203332  
DEFINITION Sequence 5 from patent US 6365361.  
ACCESSION AR203332  
VERSION AR203332.1 GI:2149694  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
Unclassified.

REFERENCE 1 (bases 1 to 1679)  
AUTHORS Taniuchi,T. and Mizukami,J.  
TITLE Method for identifying or screening agonist and antagonist to PPAR  
JOURNAL Patent: US 6365361-A 5 02-Apr-2002;  
FEATURES  
source location/Qualifiers  
1. 1679  
BASE COUNT 477 a 405 c 377 g 420 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,98e-157 Length: 1679  
Score: 2196.50 Matches: 439  
Percent Similarity: 87.15% Conservative: 2  
Best Local Similarity: 86.76% Mismatches: 2  
Query Match: 62.61% Indels: 63  
DB: Gaps: 2  
US-09-931-007a-1 (1-688) x AR203332 (1-1679)  
QY 1 MetGlyThrLeuGlnGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
Db 159 ATGGGTCAAACTGTGGGAGATTCCTATTGTACCCAGAAACGATTCTCTCATGATPACA 218  
QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40  
Db 219 CTGTCTGCAAAACATATATCACAAGATGACCATGGTTGACACAGAGATTCGCTGGCCC 278  
QY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
Db 279 ACCAACTTGGGATCAGCTCCGTGATCTCCGTAATGAGAAAGCACTCCCATCTCTTT 338  
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80  
Db 339 GATATCAAGCCCTTCACTACGTGACTTCCAGCATTTCTTACTCCACATPACGAAGAC 398  
QY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysGln 100  
Db 399 ATTCCATTACAAACAGATCCAGTGTTCAGATTTACAAAGTATGACTGTAACCTTCAA 458  
QY 101 GlnTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
Db 459 GAGTACCAAAAGTGAATCAAAAGTGAGAGCTCATCTCACCTTATTTATTCGAGAAAGCT 518  
QY 121 GlnLeuTyrAsn 124  
Db 519 CAGCTTACATTAAGCTCATGAAGAGCCTTCCAACTCCCTCATGCAATTGAATGTCTGT 578  
QY 124 124  
Db 579 GTCGTGTGAGATTAAGCTTCTGATTTCACTATGAGATTCATGCTTGTGAAGATGCAAG 638  
QY 124 124  
Db 639 GGTTCCTCCGAGAACATCAGATTGAGCTTATCTATGACAGATGTGATCTTAACGT 698  
QY 125 125  
Db 699 CGGATCCACAAAAAAGTGAATAATATGTCAGTACGTGCGTTTCAGAAATGCGTTGCA 758  
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGlnLys 158  
Db 759 GTGGGAGTGTCTCATATGCGCATGAGTTTGGCGGAGATCCGACAGGCCGAGAAAGAAAG 818  
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
Db 819 CTGTTGGCGGAGATCTCCAGATATGACACAGCTCAATCAGATCCGCTGACCTCGT 878  
QY 179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197  
Db 879 CAGGCCCTGCGCAAAACATTTGTATGACTCATACATTAAGTCTTTCGCGTGACCAAAAGCA 938  
QY 198 LysAlaArgAlaIleLeuThrGlnGlyThrThrAspLysSerProPheValIleTyrAsp 217

|||||  
Db 939 AAGGAGAGGCATCTTGCACAGAAAGACACAGCAATCCATCTGTTATCATGAC 998  
Oy 218 MetasensSerLeuMetGclYgluAspLysIleLysPheLysHisIleThrProLeuGln 237  
Db 999 ATGATTCCTTATATATATGAGACATATAATCAAGTTAAACACATCCACCCCTGCACG 1058  
Oy 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257  
Db 1059 GAGCAGAGCAAGAGAGTGGCCATCCGACATCTTTCAGAGGCTGCAGTTGCTCCGGGAG 1118  
Oy 258 AlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277  
Db 1119 GGTGGCAGAGATCATCAGAGATGCCAAAAGCATTCCTGGTTTGTAAATCTTGACTTG 1178  
Oy 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAla 297  
Db 1179 AACGACCAAGTACTCTCTCAATATATGAGATGCCAGATCATTTACACATGCTGGCC 1238  
Oy 298 SerLeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGlu 317  
Db 1239 TCCCTGATGATAAATAGAGGGGTCTCATATCCGAGGGCCAGGCTTCATGACAAAGGAG 1298  
Oy 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337  
Db 1299 TTTCTAAAGAGCCTCCGAAGCCTTTGGTGTATGAGCCCAAGTTGAGTTGCT 1358  
Oy 338 ValLysPheAsnAlaLeuGlnLeuAspSerAspSerAlaIlePheIleValIle 357  
Db 1359 GTGAAGTTCAATGCATCGTAATTAATGACACGCACTGGCAATATTTATTCCTGCTATT 1418  
Oy 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377  
Db 1419 ATTCTAGTGGAGACCGCCACAGTTTGGTGAATGTGAAGCCCATTTGAAGACATTCAGAC 1478  
Oy 378 AsnLeuLeuGlnAlaLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeu 397  
Db 1479 AACCTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAACCATCTGAGCCCTCACAGCTG 1538  
Oy 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417  
Db 1539 TTTGCCAAGCTCTCTCCAGAAATGACAGCCTCCAGACAAATGTGTACGGAAACAGCTGCG 1598  
Oy 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437  
Db 1599 CTACTGACAGGTGATCAGAAAGACGAGACAGCATGAGTCTCACCCGCTCCTGCAGAGG 1658  
Oy 438 IleTyrLysAspLeuTyr 443  
Db 1659 ATCTACAAGACTGTGTAC 1676

RESULT 7  
HUMPPARG 1679 bp mRNA linear PRI 06-FEB-1999  
LOCUS Homo sapiens mRNA for PPAR gamma2, complete cds.  
DEFINITION DB3233 DB3136  
ACCESSION D83233.1 GI:1199768  
VERSION PPAR gamma2.  
KEYWORDS PPAR gamma2.  
SOURCE Homo sapiens adipose cDNA to mRNA.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1679)  
AUTHORS Kato, S.  
TITLE PPAR gamma2  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 1679)  
AUTHORS Kato, S.  
JOURNAL Direct Submission  
TITLE Submitted (29-JAN-1996) Shigeaki Kato, Tokyo University of  
Agriculture, Department of Agricultural Chemistry; Sekuragaoka  
1-1-1, Setagaya-Ku, Tokyo 156, Japan (E-mail:xtakeem.u-tokyo.ac.jp,  
Tel:03-5477-2318, Fax:03-5477-2619)

COMMENT On or before Mar 17, 1999 this sequence version replaced  
gi:2077947, gi:1197807.  
FEATURES  
SOURCE DB3136: Submitted (18-Jan-1996).  
location/Qualifiers  
1. 1679  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/tissue\_type="adipose"  
159..1679  
/codon\_start=1  
/product="PPAR gamma2"  
/protein\_id="BA018949.1"  
/db\_xref="GI:1945622"  
/translation="MGETLGDSPIDPSDSFTDTLSANISQEMTWDFEIAFMPNPG  
ISSVDLSVMEEDSHSPDIKPTTVDFESSISTPIHEDIPTFTDPPVADYKYDLQDY  
OSAIKVPASPPYSSEKTOLYNKHSPNSLMAIECRVGSAGASGFYHVAEGCK  
GEFRRTRIKTIYRCDLNCRIKHRSNKOCYGFQKCIAGVGSNHNIRGRIAOAK  
EKLALISSDIDQNPESADIRQALNHLIDYSIKSPILTKAKARAILITKTTDSPE  
VITDMSLMGEDKIKFRHITPDLQESKEVAIRIFQCGFRSEVAEVAITEVAKSI  
FVNLDLDQVETLKYGVHEITYMLASLMKMDGLISEGQFMTREFLSLRPFQDP  
MEKFEFAVENALELSDSLAIPFAVILISGDRGLNPKPIEDIDNLLQALEQL  
KLNPRESSQLFAKLQKMTDLROIIVTEHYQLQVYIKKTTDMSLHPLQETVLDLY"  
BASE COUNT 477 a 405 c 377 g 420 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 5,98e-157 Length: 1679  
Score: 2196.50 Matches: 439  
Percent Similarity: 87.15% Conserves: 2  
Best Local Similarity: 86.76% Mismatches: 2  
Query Match: 62.61% Indels: 63  
Gaps: 2  
US-09-931-007a-1 (1-688) x HUMPPARG (1-1679)  
Oy 1 MetGlyLutThrLeuGlyAspSerProIleAspProGlnSerAspSerPheThrAspThr 20  
Db 159 ATGGGTGAAGCTCTGGGAGATTCCTATGTGACACAGAAAGCGATTCTCTCAGTATGACA 218  
Oy 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40  
Db 219 CTGTCTCAACCATATTCACAAGAAATGACCATGGTGTGACACAGAGATGCATTCGTGGGCC 278  
Oy 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
Db 279 AACCACTTTGGGATAGCTCCGTGGATCTCTCCGTAATGAAGAACATCCACATCCTTT 338  
Oy 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80  
Db 339 GATATCAAGCCCTTCACACTGCTGACTTCTCCAGCATTTCTACTCCACATTAAGAGAC 398  
Oy 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
Db 399 ATTCATTTCACAGACAGATCCAGTGGTTCAGATTAACAATGATGACCTTAACCTCAA 458  
Oy 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
Db 459 GAGTACCAAAATGCAATCAAAAGTGAGAGCTCAGCTCCACTTATTTATTCGAGAGAAGCT 518  
Oy 121 GlnLeuTyrAsn-----  
Db 519 CAGCTCTCAATTAAGCCCTCATGAAGAGCCTTCCAACTCCCTCATGGCAATTGAATGTGCT 578  
Oy 124 -----  
Db 579 GTCGTGGAGATAAAGCTTCTGATTTTCATATGAGATTCAATGCTTGTGAAGATGCAAG 638  
Oy 124 -----  
Db 639 GGTTCCTCCGAGAACACATAGATTTAGCTTATCATGACAGATGATGATCTTAAGCTGT 698  
Oy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
|||||

```

Db      699 CGGATCCACAAAAAGTAGAAATAATGTCAGTCTGCGTTTCAGAAATGCTTGCA 758
Oy      139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleGly 158
Db      759 GTGGGATGTTCTCATATATCCATCAGGTTTGGGCGGATGCGACAGGCCAGAGAGANG 818
Oy      159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg 178
Db      819 CTTGTGGCGGAGATGTCCTCCAGTATGACAGCTCAATCCAGAGTCCGCTGACCTCCGT 878
Oy      179 ---AlaLeuAlaHisIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 197
Db      879 CAGGCGCTGGCAAAACATTGTTGATGATCATATCAATAAATCCCTCCGCTGACCAAGCA 938
Oy      198 LysAlaArgAlaIleLeuThrGlyIleuThrAspIleuThrAspIleuThrAspIleuThrAsp 217
Db      939 AAGCGAGGCGCATCTTTCACAGGAAAGACACAGACAAATCCATTCCTTATCTATGAC 998
Oy      218 MetAsnSerLeuMetMetGlyGluAspIleuIleuIleuIleuIleuIleuIleuIleu 237
Db      999 ATGAATTCCTTATGATGAGAGAGATAAATCAAGTCAACATCATCACCCTGCGAG 1058
Oy      238 GluGlnSerIleuGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
Db      1059 GAGCAGAGCAAGAGAGGTCCTCCATCTTTCAGGCGCTGCGCAGTTCCGCTCGAGAG 1118
Oy      258 AlaValGlnGluIleuThrGlyIleuThrAspIleuThrAspIleuThrAspIleuThrAsp 277
Db      1119 GCTGTGCGAGAGATCAGACAGATGCGCAAAAGCATTCCTGTTTGTAAATCTTGAC 1178
Oy      278 AsnAspGlnValIleuThrLeuIleuIleuIleuIleuIleuIleuIleuIleuIleu 297
Db      1179 AAGCAGCAAGTAACTCTCTCAATATGAGATGCAAGATCATTTTACCATGCTGCGC 1238
Oy      298 SerLeuMetAsnIleuAspGlyValIleuIleuIleuIleuIleuIleuIleuIleu 317
Db      1239 TCCCTGATGATAAAGAGTGGGCTCTCATATCCGAGGCGCAAGCTTCTATGCAAGGAG 1298
Oy      318 PheLeuIleuSerLeuArgIleuProPheGlyAspPheMetGluProIleuPheGluPheAla 337
Db      1299 TTTCTAAAGAGCTGCGAAAGCCTTTTGGTGTGACTTATGAGAGCCCAAGTTTGTGCT 1358
Oy      338 ValIleuPheAsnAlaIleuIleuAspSerAspIleuAlaIlePheIleuAlaValIle 357
Db      1359 GTGAAGTTCATGACTGGAATGATGACAGCAGCTGGCAATTTTATTCCTGTCAT 1418
Oy      358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValIleuProIleuGluAspIleuGln 377
Db      1419 ATTCTCAGTGGAGACCGCCAGGTTTCTGAAATGTAACCCCATTTGAACATTTCAAGAC 1478
Oy      378 AsnLeuLeuGlnAlaIleuGluIleuGlnLeuIleuAsnHisProGluSerSerGlnLeu 397
Db      1479 AACCTGCTCAAGCCCTGGAGCTCCAGCTGGAAGCTGAACCATCTGATCTTCACAGCTG 1538
Oy      398 PheAlaIleuLeuGlnIleuIleuMetThrAspLeuArgGlnIleuAlaIleuIleuAla 417
Db      1539 TTTCGCAACCTCTCTCAGAAATGACAGACCTCAGACAGATTTGCACGGAACAGCTGACG 1598
Oy      418 LeuLeuGlnValIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 437
Db      1599 CTACGTCAGATGATCAAGAGAGAGACAGACATGAGTCTTCACCCGCTCTCTGACAGAG 1658
Oy      438 IleTyrIleuAspLeuTyr 443
Db      1659 ATCTACAGAGACTTGTAC 1676

```

```

RESULT 8
LOCUS   AY048698
DEFINITION Macaca fascicularis peroxisome proliferator-activated receptor
ACCESSION AY048698
VERSION AY048698.1 GI:21552436

```

```

KEYWORDS
SOURCE
ORGANISM
Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
Cercopithecinae; Macaca.
REFERENCE
1 (bases 1 to 1849)
Zhou,J., Wilson,K.M. and Medh,J.D.
Genetic analysis of four novel peroxisome proliferator activated
receptor-gamma splice variants in monkey macrophages
Biochem. Biophys. Res. Commun. 293 (1), 274-283 (2002)
JOURNAL
MEDLINE
22050014
PUBMED
12054596
REFERENCE
2 (bases 1 to 1849)
Zhou,J., Wilson,K.M. and Medh,J.D.
Direct Submission
Submitted (27-JUL-2001) Internal Medicine, The University of Iowa,
200 Hawkins Drive, Iowa, IA 52242, USA
FEATURES
source
1..1849
location/Qualifiers
/organism="Macaca fascicularis"
/db_xref="taxon:9541"
/cell_type="macrophage"
1..1849
/gene="PPARGgamma5"
199..1626
/gene="PPARGgamma5"
/note="transcription factor"
/codon_start=1
/product="peroxisome proliferator-activated receptor gamma
5"
/protein_id="AAL05268.1"
/db_xref="GI:21552437"
/translacion="MVDTPMPFPTNNGISVDSWDDSHSRDIPPTTVDSSIS
APHEIDIPFTPTDPMVADYRDKLQRYQAIKVPSPYSEKTLQYKPHPEPSN
SLMAIEECVGDKASGFHVGNAHCEGCKGFFRTIRLKLIDRCDLNCRHKSRNC
QYCFRQCLAVGMSHNAIRFGMPQAEKRLAEISDIDLPNSESADRLAKHLVD
SYIKSPFLTKAKARAILTKTDDKSPFVINDMSLMGDEKIKFHTITPOEORSKVA
IRIQGQFQSVFVAVOIEITERYAKIIPFVINDMDVLTLYGVHEITIMIASLMK
DGVISRGQPMFPEPLKSRKPGDPMKFEPAVKFNNLEDDSLAIFIVITIS
GDRPGLINVRFIEDIDNLDLAEFLDLKLHNPSSQLFAKLQKMTDLROIIVTEHVL
LOVIRKETDMSLHPLDLEIYKDY"
BASE COUNT 554 a 437 c 396 g 462 t
ORIGIN
Alignment Scores:
Pred. No.: 4.18e-156 Length: 1849
Score: 2186.00 Matches: 438
Percent Similarity: 82.37% Conservative: 6
Best Local Similarity: 81.26% Mismatches: 5
Query Match: 62.31% Indels: 90
DB: 9 Gaps: 2
US-09-931-007a-1 (1-688) x AY048698 (1-1849)
Oy      1 MetGlyIuThrLeuGluIAspSerProIleAspProGlnIleuSerPheThrAspThr 20
Db      25 ATGGGTGAACCTGTGGGAGATTCTCTATTGACCCAGAAAGCATTCCTTACTGACACA 84
Oy      21 LeuSerAlaAsnIleSerGln----- 27
Db      85 CTGTCTGCAAAACATATCACAGTAGAGACTGGTTTACCATGATGATGATCTCTGACG 144
Oy      28 -----GluMetThrMetVal 32
Db      145 TCGGTATCCACCACCTCGGCTCCCAAGTCTGGGATTACAAAGCAATATTCATGTT 204
Oy      33 AspThrGluMetProPheThrProThrAsnPheGlyIleSerSerValAspLeuSerVal 52
Db      205 GACACAGAGATGCCATTTTGGGCCACCACTTTGGGATCAGCTCCGIGATCTCTCCGTA 264
Oy      53 MetGluAspHisSerHisSerPheAspIleuIleuIleuIleuIleuIleuIleuIleuIleu 72

```

Db 265 ATGAGACCACTCCACCTCTTGACATCAAGCCCTTCACTGTTGACTTCCAGC 324  
 QY 115SerThrProHisTyrGluAspIleProPheThrArgThrAspProValAlaAsp 92  
 Db 325 ATTTGCGTCCACATTATGAGACATTTCCATTCAAGAACATCGATGGTGGGAT 384  
 QY 93 TyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSer 112  
 Db 385 TATTAAGTACACCTGAAACTCAAGAGTACCAAGTGCAATCAAGTGAGCGCTGCTT 444  
 QY 113 ProPheTyrSerGluTyrThrGlnLeuTyrAsn----- 124  
 Db 445 CCACCTTATTTATCTGAGAACTCACTCTACAAATTAAGCTTCACGAAAGCCCTTCCAC 504  
 QY 124 ----- 124  
 Db 505 TCCCTCATGCGAATCGAATGCTGTCTGGGAGATTAAGCTTCTGATTCACATGCA 564  
 QY 124 ----- 124  
 Db 565 GTTCATGCTTGTGAGAGATGCAAGGTTTCTCCGAGAACAAATCAGATTGAAGCTTATC 624  
 QY 125 -----ArgsnLysCysGlnTyr 130  
 Db 625 TATGACAGATGTGATCTTAACTCTCGGATCCACAAAAAGTGAATAATATGTCAGTAC 684  
 QY 131 CysArgPheGlnLysCysLeuAlaValAlaGlyMetSerHisAsnAlaIleArgPheGlyArg 150  
 Db 685 TGTGCGTTTCAGAAATGCCCTTGCAAGTGGGATGTTCATATGATCCATCAGTTTGGGCG 744  
 QY 151 MetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeu 170  
 Db 745 ATCCCAACACCCAGAGAGGAGCTGTGGCGAGATCTCCAGTATATGACACGACGCTG 804  
 QY 171 AsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLys 190  
 Db 805 AACCCAGAGTCCGCTGACCTCCGCGCCCTGGCAAAACATTTGATGATCACTATACATTAAG 864  
 QY 191 SerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrAspLys 210  
 Db 865 TCTTTCCTCCGCTGACCAAGCAAGGAGGCGCATCTTGACAGAAAGCAACAGACAA 924  
 QY 211 SerProPheValIleTyrAspMetAsnSerLeuMetMetCylLysAspLysIleLysPhe 230  
 Db 925 TCCACATTCGTTATCTATGACATGAATCTCTTATGATGGAGAAATTAATCAAGTTTC 984  
 QY 231 LysHisIleThrProLeuGlnGlnGlnSerLysGluValAlaIleArgIlePheGlnGly 250  
 Db 985 AACACATCCACCCGCTGACGAGACAGCAAGAGGTGCTATCCGCAATCTTTCAGGGC 1044  
 QY 251 CysGlnPheArgSerValGluAlaValAlaGlnGluIleThrGlyTyrAlaLysSerIlePro 270  
 Db 1045 TGGCAGTTTCGCTCGGTGGAGGCTGTGACAGAGATCACAGATATGCCAAAGCATTTCT 1104  
 QY 271 GlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGln 290  
 Db 1105 GGTGTTGTAATCTTGACTTGAACGACCAAGTAACTCTCTCAATATGAGAGCCATGAG 1164  
 QY 291 IleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlnGly 310  
 Db 1165 ATCATTTTACCAATGCTGGCTCTTGATGAATAAAGATGGGTTCCTCATATCCGAGGGC 1224  
 QY 311 GlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMet 330  
 Db 1225 CAGGCGTTCATGACAGGGAATTTTAAAGAGCTCGAAAGCCTTTGGTGACTTTATG 1284  
 QY 331 GluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAspAspSerAspLeu 350  
 Db 1285 GACCCCAAGTTTGAAGTTTGTGGAAGTTCAATGCACTGGAATTAAGTACACGAGCTTG 1344  
 QY 351 AlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLys 370  
 Db 1345 GCAAATTTATTTCTGCTCATTTATTTCACTGAGATGCGCCAGGTTTGTGAAATGTGAAG 1404

QY 371 ProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLysLeuAsn 390  
 Db 1405 CCCATTGAAGACTTTCAGACACACGCTTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAAC 1464  
 QY 391 HisProGlnSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGln 410  
 Db 1465 CACCTGAGTCCCGCAGGCTGTTTGGCAAGCTGCTCCACAAATATGACACACTCCACAG 1524  
 QY 411 IleValThrGluHisValGlnLeuLeuGlnValIleLysLysThrGluThrAspMetSer 430  
 Db 1525 ATTTGTCAGGACACGCTGACGCTACTGCAAGTGATCAAGAGACGAGACAGATGAGCT 1584  
 QY 431 LeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyrAlaThrAlaIleLeuThr 449  
 Db 1585 CTTCACCCGCTCTGACAGGAGATCTCAAGAGACTTGTACTAGCAGAGAGCTCCGAGC 1641  
 RESULT 9  
 AF059245 1758 bp mRNA linear NAM 21-APR-1998  
 LOCUS  
 DEFINITION  
 (PPARG) mRNA, complete cds.  
 ACCESSION  
 AF059245  
 VERSION  
 AF059245.1 GI:3065956  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Sus scrofa.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE  
 1 (bases 1 to 1758)  
 Houseknecht, K.L., Blawell, C.A., Portocarreiro, C.P., Spurlock, M.E.  
 and Heady, A.  
 Cloning and regulation of porcine peroxisome proliferator-activated  
 receptor gamma  
 TITLE  
 JOURNAL  
 Unpublished  
 2 (bases 1 to 1758)  
 Houseknecht, K.L., Blawell, C.A., Portocarreiro, C.P., Spurlock, M.E.  
 and Heady, A.  
 Direct Submission  
 Submitted (13-APR-1998) Animal Sciences, Purdue University, 1026  
 Poultry Science Building, West Lafayette, IN 47906-1026, USA  
 FEATURES  
 source  
 1..1758  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 1..1758  
 /gene="PPARG"  
 17..1531  
 /gene="PPARG"  
 /codon\_start=1  
 /product="peroxisome proliferator-activated receptor gamma  
 2"  
 /protein\_id="AAC1348.1"  
 /db\_xref="GI:3065957"  
 /translation="MGFTLGDLSLDIPESDARFDLTLSANISQEVTVNDTEMPMPNFGI  
 SSVDLSVMDHSHSFEDIKFTTYDFFSSSTFHYEDIPRPADPMADYKYLKQDYQ  
 SAIKVEPSPYSEKTOLYNKHPHEPNSNLATCRCGKASGFHYGAGCGCG  
 FFRRTIRLKLTYDRCDLNCRIHKRSRNCQVCRQGLKAVGMSHNAIFEGMPARE  
 KILALISDIDOLNPESADYRLAKHLIDYSIKSPRLTKAKARALITGKTDKSPYI  
 YDMSLKMGEDEIKRHTPLQDSKEVAIRFQGCOPRSYEAQVETETAKNIPGV  
 NDLNDQVTLKRYGVHETITYMLASLMMKDDVLISEGGFMYREFLKLSPFDPMF  
 PKFEFVAENMLEDDSDLAIFIAVILISGRPGILNPKPIEDIDONLQALELQKL  
 NHPESQAFKALDKRMFLDRIIVTEHVOLLQVIKKETDMSLHPLQETRYKDY"  
 BASE COUNT 522 a 416 c 390 g 430 t  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,1e-155 Length: 1758  
 Score: 2172.50 Matches: 440  
 Percent Similarity: 82.58% Conservative: 15  
 Best Local Similarity: 79.85% Mismatches: 25  
 Query Match: 61.93% Indels: 71  
 DB: 4 Gaps: 4

```

US-09-931-007a-1 (1-688) x AF059245 (1-1758)
OY 1 MetGlyGluThrLeuGlyAspSerProIleasProGluSerAspSerPheThrAspThr 20
    |||||
Db 17 ATGGGGAAACTCTGGGAGATTCTTATTTGACCCAGAAAGCGATCCCTTC--GACAG 73
OY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
    |||||
Db 74 CTGTCTGGCAACATTTCACAGAGGTGACCATGTTGACCCGAGATGCCGTTTGGGCC 133
OY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
    |||||
Db 134 ACCAACTTTGGATCAGCTCTGTGACCTGTGGTGATGAGACACCTCCACCTCCTTT 193
OY 61 AspIleIysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
    |||||
Db 194 GACATCAACCCCTTCACCACTGTGATTTCTCCACATTTCACATTCACATCAAGAGAC 253
OY 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGln 100
    |||||
Db 254 ATCCATTTCCGAGAGCTGATCCATGTTGACAGATTATTAAGTACGACCTGAGCTCCAG 313
OY 101 GlnTyrGlnSerAlaIleIysValGluProAlaSerProProTyrTyrSerGluIleThr 120
    |||||
Db 314 GACTACCAAGAGTCCATCAAGTGAAGCCTGTGTCGCCCTTATTTCTGAAAAAGACT 373
OY 121 GlnLeuTyrAsn----- 124
    |||||
Db 374 CAGCTGTACAAACAAACCTCAGAGAGCCCTTCCAACTCCCTCATGCAATTGATGCCGT 433
OY 124----- 124
Db 434 GTCGTGGGGATAAAGCCTCGGGTTCACATGAGATTGATGCTTGTGAAGATGCAAG 493
OY 124----- 124
Db 494 GGTTCCTCCGAGAGACTATCAGATTGAAGCTTATTTATGATAGTGGGATCTTAAGT 553
OY 125-----ArgAsnIysCysGlnTyrCysAlaGlyPheGlnIleCysLeuAla 138
    |||||
Db 554 CGGATCCACAAAAAAGAAATGAATGAATGCTCAGTACTGCTGCTTTCACAAAAAGCCTCT 613
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleGln 158
    |||||
Db 614 GTGGGGATGTCTCATACGCCATCAGGTTTGGCGGATGCCACAGGCTGAGAGAGAGAG 673
OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
    |||||
Db 674 CTGTGGCAGAGATCTCCAGTATGACGACGACGAGCTGAACCCAGAGCTGTGACCTCCGC 733
OY 179 AlaLeuAlaIleHisLeuTyrAspSerTyrIleIysSerPheProLeuThrIleAlaIys 198
    |||||
Db 734 GCGCTGGCAAAACCTGTTATGACATCAACAATAAAGTCCCTCCCTGACCAAAACAAG 793
OY 199 AlaArgAlaIleLeuThrGlyIleThrThrAspIleAspPheSerProPheValIleTyrAspMet 218
    |||||
Db 794 GCGAGGGCGATTTTACAGAGAAAGACACAGACAAATACACCTTTGTATCTATGACATG 853
OY 219 AsnSerLeuMetMetGlyIleAspIleIysPheIleIysPheIleThrProLeuGlnIle 238
    |||||
Db 854 AATTCCTTAATGATGGAGAGATAAATCAAGTCAAAACATCAACCCCTCGCAGAGAG 913
OY 239 GlnSerIysGlnValAlaIleArgIlePheGlnIleGlyCysGlnPheArgSerValGluAla 258
    |||||
Db 914 CAGACCAAAAGAGTGGCCATTCGATCTTTCAGGGGTCTCAATTTGCGCTCGTAGAGCC 973
OY 259 ValGlnGluIleThrGluTyrAlaIysSerIleProGlyPheValAsnLeuAspLeuAsn 278
    |||||
Db 974 GTGCAGAGATACAGAGATGCCAAGAACAATCCTCGGTTTGTAAACCTTGACCTGAAT 1033
OY 279 AspGlnValThrLeuLeuIysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
    |||||

```

```

Db 1034 GACCAGTAACCTTCCTAAGATGAGCTCCACAGATCATTTACACCATGCTGGCTCC 1093
OY 299 LeuMetAsnIysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPhe 318
    |||||
Db 1094 TTGATGAATTAAGAGCGGGTCTCATCTCCAGAGGGCCAAAGATTCATACAAAGGAGTTT 1153
OY 319 LeuIysSerLeuArgIysProPheGlyAspPheMetGluProIysPheGluPheAlaVal 338
    |||||
Db 1154 CTCAGAGACCTTGAGAAACCCCTTTGGTACTTTATGAGGCCCAAGTTGAGATTGGCTGTG 1213
OY 339 LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
    |||||
Db 1214 AAGTTCAACGACCTGAGATTTAGATGACAGCGACCTGGCGATATTATATGCTGATATT 1273
OY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIysProIleGluAspIleGlnAspAsn 378
    |||||
Db 1274 CTCAGTGGAGACCCGCCAGGTTTGTGAATGTGAAGCCCATTCGAGGACATTCAGACAT 1333
OY 379 LeuLeuGlnAlaLeuGluLeuGlnIleuIysLeuAsnHisProGluSerSerGlnLeuPhe 398
    |||||
Db 1334 TTGCTGCAAGCCTTGAGACTGCTGAGTCAAGTTGAACCAACCCAGAGTCTCCACAGCTTT 1393
OY 399 AlaIysLeuLeuGlnIleIysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
    |||||
Db 1394 GCTAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTGACAGACACCTGCAGCTG 1453
OY 419 LeuGlnValIleIysIysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
    |||||
Db 1454 CTGCAAGTAATTAAGAAAGACAGAGACGACATGACCTTCACCCACTCTCAGAGAAATA 1513
OY 439 TyrIysAspLeuTyrAlaThrAlaIleLeuThrGlyIleThrThrAspIleSerPro--- 457
    |||||
Db 1514 TACAAAGACTGTGAC-----TACCAAAAGATCCACGTTCCACCGCA 1555
OY 458 --PheValIleTyrAspMetLeuAsnSerLeuMetGlyGluAspIysIleIysPheIys 476
    |||||
Db 1556 ACATCTCCTCTCTCCATTTGACATATATTTTGGGGGAAAAAATGTGACACTAATAA 1615
OY 477 HisIleThrProLeuGlnGluGlnSerIysGlu 487
    |||||
Db 1616 AAAATTAAGTGAAGAAAGCGTTTGAAAAAGAA 1648

```

RESULT 10  
 AF103946  
 LOCUS 1539 bp mRNA linear MAM 12-Apr-2000  
 DEFINITION Sus scrofa peroxisome proliferator activated receptor gamma 2 (PPAR  
 gamma 2) mRNA, complete cds.  
 ACCESSION AF103946  
 VERSION AF103946.1 GI:4378476  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Sus scrofa.  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 REFERENCE  
 1 (bases 1 to 1539)  
 DING S.T., McNEEL R.L. and MERMANN, H.J.  
 Expression of porcine adipocyte transcripts: tissue distribution  
 and differentiation in vitro and in vivo  
 Comp. Biochem. Physiol. B, Biochem. Mol. Biol. 123 (3), 307-318  
 (1999)  
 MEDLINE 99410873  
 PUBMED 10481259  
 REFERENCE 2 (bases 1 to 1539)  
 DING, S.T. and MERMANN, H.J.  
 Direct Submission  
 Submitted (03-NOV-1998) Pediatrics, Baylor College of Medicine,  
 1100 Bates St., Houston, TX 77030, USA  
 FEATURES  
 source  
 1..1539  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 1..1539  
 /gene="PPAR gamma 2"



CDS  
16. .1530  
/gene="PPAR gamma 2"  
/codon\_start=1  
/product="peroxisome proliferator activated receptor gamma 2"  
/protein\_id="MAD19577.1"  
/db\_xref="GI:4378477"  
/translation="MGELGDSLIDPESDAFDPLSANISOEVNMDTEMPFWPTNGI  
SSVDLSVMDHSHSPDIKPFYTVDESSISTPHYEDIPEPRADPMVADYKYDLKDYQ  
SAIKVEPVPPYSEKTOYLNKPHEPSSSLAIEKRVCGKSGASGFVHAGCEGKG  
FFRRTRLRKLIYDRDLNCRIRHKSRNKQCYRFOCKLAIVGSHNAREGRPOAEKE  
KILAEISDIDLNPSADRLRLAKHLDSYIKSPFLTAARAILITGTTTKSPRY  
YKNSLDMKEDKIKRHHITPQEOSKEVAIRIFQOGQFSEVAODETTEYANIDPEV  
NLDELNDQVTLKYGVAHEIITYMLASLMNDGVLISGQGFMTREFLSLRKPDGFM  
PKFEFAVKENALELDDSDLAIFIAVILISGDRPGLNPKVPIEDIDNLQALELQKL  
NHPSOLEFAKLQKMTDLRQIIVTEHVOLLQVIKKTETMSLHPLOEITYKDYL"

BASE COUNT 421 a 384 c 369 g 365 t  
ORIGIN

Alignment Scores:  
Pred. No.: 4.14e-155 Length: 1539  
Score: 2171.50 Matches: 432  
Percent Similarity: 86.73% Conservative: 6  
Best Local Similarity: 85.54% Mismatches: 4  
Query Match: 61.90% Indels: 63  
DB: Gaps: 2

US-09-931-007a-1 (1-688) x AF103946 (1-1539)

QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20  
16 ATGGGGAACCTCTGGAGATCTCTATTGACCCAGAAAGCATGCTTC---GACAG 72  
Db  
QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThr 40  
73 CTCTCGCAACATTTCCACAAGAGGTGACCATGTGACCCAGATGCCGTTTGGCCC 132  
Db  
QY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
133 ACCAACTTTGGGATCAGCTGTGTGACCTGTGATGAGCAGACACTCCACTCTTT 192  
Db  
QY 61 AspIleLeuSerProPheThrThrValAspPheSerSerIleSerThrProHisThrGluAsp 80  
193 GACATCAAGCCCTTCACACCTGTGATTTCTCCAGATTTCACCTCACATATGAAGAC 252  
Db  
QY 81 IleProPheThrArgThrAspProValIleAlaAspTyrIleTyrAspLeuLysLeuGln 100  
253 ATCCCATTCGCCGAGAGCTGATCCATGTTGCGAGATTATAGACAGCCTGGAAGCTCCAG 312  
Db  
QY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
313 GACTACCAAGTCCATCAAGAGGAGCGGTGTCCCGCTTATTTATCTGAAAAGACT 372  
Db  
QY 121 GlnLeuTyrAsn----- 124  
373 CAGCTGTACAAACCTCAGAGAGCCTTCCAACTCCCTCATGGCAATTGAATCCCGT 432  
Db  
QY 124 ----- 124  
433 GTCTGTGGGATAAAGCCTCGGGGTTCCCATATGAGATTCACTTGTGAGAGATGCAG 492  
Db  
QY 124 ----- 124  
493 GGTTCCTCCGAGACTATCAGATTGAAGCTATTATGATGAGTGCATCTTAAGTGT 552  
Db  
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
553 CGGATCCACAAAAAAGTGAATAATATGCACTACTGCGTTTGAAGAAATCCCTGCT 612  
Db  
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnIleGlnLysGluLys 158  
613 GTGGGATGTCTCATACGCCATCAGGTTTGGCGGATGCCACAGGCTGAGAGGAGAG 672  
Db

QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
673 CTGTGGCAGAGATCTCCAGTGTATTCAGACACCTGAACCCAGACGCTGTGACTCCGC 732  
Db  
QY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198  
733 GCCCTGGCAAGCATTTGATGACTCATCAATAATCTCTCCCGGTGACCAAGCAAG 792  
Db  
QY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
793 GCGAGGCGCATCTTGACAGAGAAAGACCACAGACAAATCACCTTTTCTATGTGACATG 852  
Db  
QY 219 AsnSerLeuMetLeuGlyLysPheLysIleLysPheLysHisIleThrProLeuGlnGlu 238  
853 AATTCTTATGATGGAGAGAGATATAATCAAGTTCAACACATCACCCCTGCAGAG 912  
Db  
QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258  
913 CAGAGCAAGAGGTGGCATTCGATCTTTCAGGGGTGTCACTTGTGTCCTGGAGGCG 972  
Db  
QY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
973 GTCCAGAGATCTACAGAGTATGCCAAGACATCCCTGGGTTTCTAATACCTTGACCTGAT 1032  
Db  
QY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
1033 GACCAAGTAACCTCTCTAAGTATGGCGTCCACAGATCATTTACACCATGTCGCTCC 1092  
Db  
QY 299 LeuMetAsnLysAspGlyValIleLeuIleSerGlnGlyCysGlnPheMetThrArgGluPhe 318  
1093 TTGATGAATTAAGACGGGGCTCTCATCTCCGAGGGCCAAGATTCATGCAAGGGAGTTT 1152  
Db  
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338  
1153 CTCGAAGCCTGGAAGAGCCCTTGGTGTACTTATGAGACCCAAAGTTCAGTTGCTGTG 1212  
Db  
QY 339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
1213 AAGTTCAACGACATGGAAATAGATGACAGCAGCCTGGCATATTATACCTGTCATTAAT 1272  
Db  
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
1273 CTGAGTGGAGACCGCCAGCTTGTGAAATGTCAACCCATGAGGACATTCGAAGCAAT 1332  
Db  
QY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerGlnLeuPhe 398  
1333 TTGCTCTCAAGCCTTGGAGCTGACGCTCAAGTTGAACACCCAGAGTCTCCACGCTCTT 1392  
Db  
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418  
1393 GCTAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTGACAGAGCAGCTGCAGCTG 1452  
Db  
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438  
1453 CTGCAAGTAAATAAGAAAGACAGAGAGGAGACTTACCCACTCTTACAGAGAAATA 1512  
Db  
QY 439 TyrLysAspLeuTyr 443  
1513 TACAAGACTTGTAC 1527  
Db

RESULT 11  
SSC6757  
LOCUS SSC6757 1640 bp mRNA linear MM 24-SEP-1998  
DEFINITION Sus scrofa mRNA for peroxisome proliferator-activated receptor  
gamma 2.  
ACCESSION AJ006757  
VERSION AJ006757.1 GI:3646469  
KEYWORDS peroxisome proliferator-activated receptor gamma 2; PPARg2 gene.  
SOURCE Sus scrofa.  
ORGANISM Sus scrofa.  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
REFERENCE 1 (bases 1 to 1640)

AUTHORS Grindflek, E.  
 TITLE Direct Submission  
 JOURNAL Submitted (12-JUN-1998) Grindflek E., Animal Science, Agricultural University of Norway, P.O. Box. 5025, 1432 AS, NORWAY  
 REFERENCE 2 (bases 1 to 1640)  
 AUTHORS Grindflek, E., Sundvold, H., Klungland, H. and Lien, S.  
 TITLE Characterisation of porcine peroxisome proliferator-activated receptors gamma 1 and gamma 2: detection of breed and age differences in gene expression  
 JOURNAL Biochem. Res. Commun. 249 (3), 713-718 (1998)  
 MEDLINE 98401156  
 PUBMED 9731203

FEATURES  
 source Location/Qualifiers  
 1..1640  
 /organism="Sus scrofa"  
 /db\_xref="taxon:9823"  
 /cell\_type="adipocyte"  
 61..1575  
 /gene="PPARG2"  
 61..1575  
 /gene="PPARG2"  
 /function="adipocyte differentiation, fat deposition, regulation of adipogenesis"  
 /codon\_start=1  
 /product="peroxisome proliferator-activated receptor-gamma 2"  
 /protein\_id="CA07225.1"  
 /db\_xref="GI:3646470"  
 /db\_xref="SPTREMBL:O77815"  
 /translation="MGETLGSLLIDPESDAFDLTSANI SOEVTWDTMEPWPNTFCI SSVLSVMDHSHFEDIKPPTVDPSISIPHYEDIPFPADPVAADYKVDLKQDYQ SAIKVEVSPPYSEKQOLYNKPHPEPSNSLMAIECRVCGDKASGFHYGVAECGKG FFRRTIRKLILYDRCDLNCRIHKSRNKCOCYRQKCLAVGSHNAIRFGMPAEKE KLAETISDDILNPESADLRALKHLYDSYIKRSPILKAKARAILITGKTDKSPVY YDMSLSMGEDEKIKFKHTTPLEQOSKEVAIRIPGCCFRSENAVDEITFAKNTPGY NUDLNDVTLTKYVHEIITITMLASLMKMDVLSISGQGEWTEFLSLRKPREDME KPEFAVKFNALDEDSDLAFLAVIILSGRPRLNVKRPLEDIQDMLLALLETQLK NHPSSOLFKLLQKMTDLROIYVHEVOLQVIKETEDMSLHPLQEIYKDLT"

BASE COUNT 459 a 408 c 382 g 391 t  
 ORIGIN

Alignment Scores:  
 Pred. No.: 3.03e-154 Length: 1640  
 Score: 2160.50 Matches: 430  
 Percent Similarity: 86.53% Conservative: 7  
 Best Local Similarity: 85.15% Mismatches: 5  
 Query Match: 61.59% Indels: 63  
 DB: Gaps: 2

US-09-931-007A-1 (1-688) x SSC6757 (1-1640)

Oy 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
 |||  
 Db 61 ATGGGTGAACCTCTGGGAGATTCTCTTATTGACCCAGAAAGGATGCTTC--GACACG 117

Oy 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40  
 |||  
 Db 118 CTGCTGGAACATTTCACAGAGGTGACATGCTTGACACCGAATGCCGTTTGCGCC 177

Oy 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
 |||  
 Db 178 ACCAAGCTTGGATCAGCTGTGTGACCTGCGGATGAGGACGACCACTCCACCTCTT 237

Oy 61 AspIleIleSerPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80  
 |||  
 Db 238 GACATCAAGCCCTTACACAGCTGTGATTCACGACATTCACCTCCACACTATGAAGAC 297

Oy 81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGlu 100  
 |||  
 Db 298 ATCCCATTCGCGAGAGCGATGCAATGTTGAGATTAAAGTACGACCTGAAGCTCCAG 357

Oy 101 GluTyrGlnSerAlaIleIleValGluProAlaSerProProTyrTyrSerGluIleThr 120  
 ::::|  
 ::::|

Db 358 GACTACCAAGATGCCATCAAAAGTGAGCGCTGTGTCCCGCCTTATTATTCAAAAAGACT 417  
 Oy 121 GluLeuTyrAsn----- 124  
 |||  
 Db 418 CAGCTGTACAAACACCTCAGCAAGACCTTCCAACTCCCTCATGGCAATTGATGCCGT 477

Oy 124 ----- 124

Db 478 GTCTGGGGATAAAGCCCGGGTTCCACTATGAGATTGATGTTGAAGATGCAAG 537  
 Oy 124 ----- 124

Db 538 GGTTCCTTCGAGAGCATCATGATTGAGCTTATTATGATGAGTGGACCTTAAGTCT 597  
 Oy 125 -----ArgAsnIleCysGlnTyrCysArgPheGlnIleCysIleVal 138  
 |||  
 Db 598 CGGATCCCAAAAAAGATGAATTAATGATGATGCTCGGTTCACAAATGCTCTCT 657

Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIleGlu 158  
 |||  
 Db 658 GTGGGATGTCATCAACGCCATCAGTTTGGCGGATGCCACAGAGTGGAGAGAGAG 717

Oy 159 LeuLeuAlaGluIleSerSerAspIleAspGluLeuAsnProGluSerAlaAspLeuArg 178  
 |||  
 Db 718 CTGTTGGCAGAGATCTCCAGTATGATGACACAGCTGAACCCAGAGATCTGTGACCTCCG 777

Oy 179 AlaLeuAlaIleHisLeuTyrAspSerTyrIleIleCysSerPheProLeuThrIleVal 198  
 |||  
 Db 778 GCCCTGGCAAGACCTTGTATGACATCAATAAAGTCTTCCCTCCATCAAAAGCAAG 837

Oy 199 AlaArgAlaIleLeuThrGlyIleTyrThrThrAspIleSerProPheValIleTyrAspMet 218  
 |||  
 Db 898 GCGAGGGGATCTTACAGCAAGCAACACAGCAATACACCTTTGTATCATATGACATG 897

Oy 219 AsnSerLeuMetMetGlyGluAspIleIleCysPheIleHisIleThrProLeuGlnIle 238  
 |||  
 Db 898 AATTCCTTAAGATGGAGAGATAAATCAAGTCAAAACATCAACCCCTCGAGAGAG 957

Oy 239 GlnSerIleGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluIle 258  
 |||  
 Db 958 CAGACCAAGAGGTGGCCATTCGCAATCTTTCAGGGGTGTGATTCGCTCGTGGAGGCC 1017

Oy 259 ValGlnGluIleThrGluTyrAlaIleCysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
 |||  
 Db 1018 GTGCAGAGATCACAGAGATATGCCAAGAACATCCCTGGCTTTGTTAACTTGACCTGAT 1077

Oy 279 AspGlnValIleThrLeuLeuIleTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
 |||  
 Db 1078 GACCAAGTAACTCTCTAAAGTATGGCTCCACGAGATCATTTACCATCTGCTGCCCTCC 1137

Oy 299 LeuMetAsnIleAspGlyValLeuIleSerGlnGlyGlnIlePheMetThrArgGluPhe 318  
 |||  
 Db 1138 TTGATGAATTAAGACGGGGTCTCTATCTCCAGGGCCCAAGGATTCATGACAAAGGAGTTT 1197

Oy 319 LeuIleSerLeuArgIleProPheGlyAspPheMetGluProIleCysPheGluPheAlaVal 338  
 |||  
 Db 1198 CTAAGAGACCTCGAAGAAACCTTTGGTGTACTTATGAGGCCCAAGTTCGATTTGCTGTG 1257

Oy 339 LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
 |||  
 Db 1258 AAGTTCACAGCGACATGATGATGACAGCACTGGCGGATATTATAGCTCATATTATT 1317

Oy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIleLysProIleGluAspIleGluAspAsn 378  
 |||  
 Db 1318 CTCAGTGGAGACCGCCACGTTTGTGTGATGTGAAGCCCATGAGAGACATTCAGACACAT 1377

Oy 379 LeuLeuGlnAlaLeuGluLeuGlnLeuIleCysLeuAsnHisProGluSerSerGlnLeuPhe 398  
 |||  
 Db 1378 TTGCGTCAGAGCTTGGAGCTGACGTCAAGTTGAACCAACCCAGAGTCTCCACACTCTTT 1437

Oy 399 AlaIleLeuLeuGlnIleMetThrAspLeuArgGlnIleValIleThrGluHisValGlnLeu 418  
 |||  
 Db 1438 GCTAAGCTGTCTCCAGAAATGACAGACCTCAGACAGATTGTGACAGAGACGTCAGCTG 1497

419 LeuGlnValIleIleLysLysThrGluThrAspMetSerLeuHisProLeuGlnGluIle 438  
|||||  
Db 1498 CTGCAAGTATATAAGACAGACAGACGACATGACCTTACACCTCTCTACAGGAATA 1557  
OY 439 TyrLysAspLeuTyr 443  
|||||  
Db 1558 TACAAGACTTCTAC 1572

RESULT 12  
BC021798 1782 bp mRNA linear ROD 07-AUG-2002  
LOCUS BC021798  
DEFINITION Mus musculus, peroxisome proliferator activated receptor gamma,  
clone MGC:18439 IMAGE:4191539, mRNA, complete cds.  
ACCESSION BC021798  
VERSION BC021798.1 GI:18255315  
KEYWORDS MGC.  
SOURCE house mouse.  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 1782)  
REFERENCE 1  
AUTHORS Strausberg, R.  
TITLE Direct Submission  
JOURNAL Submitted (18-JAN-2002) National Institutes of Health, Mammalian  
Gene Collection (MGC), Cancer Genomics Office, National Cancer  
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,  
USA  
NIH-MGC Project URL: <http://mgc.nci.nih.gov>  
REMARK Contact: MGC help desk  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Jeffrey E. Green, M.D.  
CDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Baylor College of Medicine Human Genome  
Sequencing Center  
Center code: BCM-HGSC  
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>  
Contact: [amgdbcm.tmc.edu](mailto:amgdbcm.tmc.edu)  
Guanarino, P.H., Garcia, A.M., Lu, X., Hu, Y.K., S.W., Hale, S.M.,  
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,  
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found  
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Series: IRAX Plate: 24 Row: i Column: 13  
This clone was selected for full length sequencing because it  
passed the following selection criteria: matched mRNA gi: 6755137.  
FEATURES  
source  
1..1782  
Location/Qualifiers  
/organism="Mus musculus"  
/db\_xref="taxon:10090"  
/map="FVB/N"  
/clone="MGC:18439 IMAGE:4191539"  
/tissue\_type="Salivary gland, 10 week old female mouse"  
/clone\_1lb="NCI CGAP\_SG2"  
/lab\_host="DH10B"  
/note="vector: pCMV-SPORT6"  
46..1563  
/codon\_start=1  
/product="peroxisome proliferator activated receptor  
gamma"  
/protein\_id="AAH21798.1"  
/db\_xref="GI:18255316"  
/db\_xref="locusID:19016"  
/translation="MGETLGDSPVDEHGAFAADALPMSTSOEITWDTMPFWPTNFG  
ISSVDSLMEHSHSFEDIKPTTVDFSSISAPHEDIPTFRADPMVADYKDLKIOEY  
QSAIKVEPASPYPSEKTLNRPHEESNLMAIECVCDKASGFAYGVYAGEGCK  
GEFRRTIRKLIDRCDLNCRIRHKSRKRCYCPROCKLAAGSMNATRFRCMOAOK  
EKLAETSSDIDOLNPESADRLAKLHYDSYISFSPKTRKAAIILTKRTDMSPV  
IYDMSLIMGEDKIKFKHITPLOSKEVAIRITOGCOFRSEVAOELTERAKINPGF  
INLNDQVTLTKYGVHEITITMLASLNNKQGVLSIEQGMTEPEFLKSLRPGDFM  
EPKPEFAVKFNALEIDSDLAIFLAVIILSGDRQLNVRKPIEDIDQNLQALQLK

LNHPESOLFARKVLQKMTDLROIUTEHVOLLHVIKKTEPDMSLHPLLQEIYKDY"  
BASE COUNT 533 a 416 c 386 g 447 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 7,33e-154 Length: 1782  
Score: 2156.00 Matches: 425  
Percent Similarity: 85.74% Conservative: 8  
Best Local Similarity: 84.16% Mismatches: 10  
Query Match: 61.46% Indels: 62  
DB: 10 Gaps: 1  
US-09-931-007A-1 (1-688) x BC021798 (1-1782)  
OY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
|||||  
Db 46 ATGGGGAACCTGGGAATTCCTCTGTGACCAAGACATGTCCTTCCTATGCA 105  
OY 21 LeuSerAlaAsnIleSerGlnIleuMetThrMetValAspThrGluMetProPheThrPro 40  
|||  
Db 106 CTGCCTATGACGACCTTCACAAAGAAATTAACATGTTGACACAGACATGCTTGGGCC 165  
OY 41 ThrAspPheGlyTleSerSerValAspLeuSerValMetGluAspHisSerHisPhe 60  
|||||  
Db 166 ACCAATCTCGAATGACGCTGTGACACCTCCGATGGAAGACACATCGCATTCCTTT 225  
OY 61 AspIleLysProPheThrValAspPheSerSerIleSerThrProHisThrGluAsp 80  
|||||  
Db 226 GACATCAAGCCCTTACCACTGACATGATTTCTCAGACATTTGCTCCACACTATGAAGAC 285  
OY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
|||||  
Db 286 ATTCATTTACAAAGAGCTACCCATGCTGCTGATTAACAAATATGACCTGAAGCTCCA 345  
OY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
|||||  
Db 346 GAATACCAAGTCGATCAAGATAGAACCTGCATCTCCACTTATATTCGTAAAAAGACC 405  
OY 121 GlnLeuTyrAsp 124  
|||||  
Db 406 CAGCTCTACAAAGCGCTCATGAAGAACCTTCACTCCCTCATGGCCATTGATGCCGA 465  
OY 124 124  
Db 466 GTCTGTGGGGAATAAAGCATCAGGCTTCCACTATGAGATTCATGCTTGAAGATGCAAG 525  
OY 124 124  
Db 526 GGTTTTTCGGAAGACCATCCGATTGAAGCTTATTTATGATAGGTGATCTTAACATGC 585  
OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
Db 586 CGGATCCACAAAAAAGTAGAATAATGAATGTCAGTCTGCTGTTTCAAGAGTCCCTCTCT 645  
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158  
|||||  
Db 646 GTGGGATGTCACAAATGCCATCAGGTTTGGCGGATGCCACAGCCGAGAGAGAGAAAG 705  
OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
|||||  
Db 706 CTGTTGGCGGAATCTCCAGTATATTCACACGAGAACCCAGAGATCTGCTATCTGCGCA 765  
OY 179 AlaLeuAlaLysIleLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198  
|||||  
Db 766 GCCCTGGCAACCATTTGATGATCATCATCAATGAATGCTCCCTGACCAAAACCAAG 825  
OY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
|||||  
Db 826 GCGAAGGCGATTTTACAGAGGAAGAACACGCAATATCACCATTGTCTATCAGACATG 885  
OY 219 AsnSerLeuMetCysGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238  
|||||  
Db 886 AATTCCTTAATGATGGAGAGATAAATCAAGTCAAAACATATACACCCCTCGAGAG 945

OY	239	glnserLysgIuValAlAlAlleqrlglebpehngIngySgInPheargSerAlgluAla	258
Db	946	CAGACCAAGAGAGGTGGCCATCCGAATTTTTCAGAGGTGGCAGATTTCGATCCGTAGAGACC	1005
OY	259	ValGIngluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuaspLeuasn	278
Db	1006	GGCCAAAGAGATCACAGAGATGCCAAAATAATCCCTGGTTTCATTAACTTATTTGCAAT	1065
OY	279	AspGInValThrLeuLeuLysTyrGlyAlaHisGluIleIleTyrThrMetLeuAlaser	298
Db	1066	GACCAAGTACTCTCTCAGATGTGGTCCATGTAGATCATCTACACGATGCTGGCTTCC	1125
OY	299	LeuMetAsnLysAspGlyValLeuIleSerGluGlyGInGlyPheMetThrArgIuphe	318
Db	1126	CTGATGAATTAAGATGAGATCTCTCATCTCAGAGGCCAAGATTAATGACAGGGAATTTC	1185
OY	319	LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal	338
Db	1186	CTCAAAACCTTCGGGAACCCCTTTGGTGACTTATGAGACCCTAAGTTGAGTTTCTGTGG	1245
OY	339	LysPheAsnAlaLeuGluLeuAspSperSerAspLeuAlaIlePheIleAlaValIleIle	358
Db	1246	AAGTTCATATGCCTGGATTTAGATGACAGTGCCTGGCTAATATTAATGCTCTCATTTAT	1305
OY	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn	378
Db	1306	CTCAGTGGAGACCGCCGACGGCTGTGTGAACGTGAAGCCCATGAGACATCCAAAGACAC	1365
OY	379	LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe	398
Db	1366	CTGCTGCAGGCGCTTGGAACCTGCAGCTCAGCGTAATCCAGACTCCTCTCAGCTGTTC	1425
OY	399	AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu	418
Db	1426	GCCAAAGTGCTCCAGAAATGACAGACTCTCAGGCGAGATGTCACAGACAGCAGCTGA	1485
OY	419	LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnIle	438
Db	1486	CTGCATGTGATCAAGAACACAGAGACAGATGAGCTTCACCCCTGCTCCAGAGATTC	1545
OY	439	TyrLysAspLeuTyr 443	
Db	1546	TACAGAGACTGTGAT 1560	
RESULT 13			
LOCUS	MMU09138	1780 bp	linear ROD 18-FEB-1995
DEFINITION	Mus musculus peroxisome proliferator activated protein-gamma-2		
ACCESSION	U09138		
VERSION	U09138.1		
KEYWORDS	GI:500639		
SOURCE	Mus musculus.		
ORGANISM	Mus musculus.		
REFERENCE	1. Mus musculus. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 1780)		
AUTHORS	Tontonoz, P., Hu, E., Graves, R.A., Budavari, A.I., and Spiegelman, B.M.		
TITLE	MPAR gamma 2: tissue-specific regulator of an adipocyte enhancer		
JOURNAL	Genes Dev. 8 (10), 1224-1234 (1994)		
MEDLINE	95011536		
PubMed	7926726		
REFERENCE	2 (bases 1 to 1780)		
AUTHORS	Hu, E.		
TITLE	Direct Submission		
JOURNAL	Submitted (21-APR-1994) Erding Hu, Cell and Molecular Biology, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA		
FEATURES	Location/Qualifiers		
Source	1..1780		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		

CDS

```

/cell_line="3T3-F442A"
/tissue_type="adipose tissue"
/clone_lib="lamdba zapir library"
40..1557
/codon_start=1
/product="peroxisome proliferator activated
protein-gamma-2"
/protein_id="AA06227.1"
/db_xref="gi:500640"
/translation="MGETLGDSPVDPHGAFADALPMSQETIMVDTEMPEMPTNG
ISSVDLSVMEDHSFPIKPTPTVDPSISAPHEVDIPTRADPMVADYKYDLQLEY
OSAIKIVEPASPSPYSSEKTOLYNRPHEPNSNIM IERNRGDASGRHYCVHACECK
GFFPRTITFLKTIYRCDLNCRIHKSSRNKOYCRFOCLAVGKSHNAIRGRNPOAK
EKLIAETISDDQLNPSADRLAKALYDSYIKSPFLYAKARAILGTGTTDSPEV
IYDNKNSLMGDEDIKFNHITPLOSKEVAIRFQGOQFSEVAVOITETAKNIDGF
INLDNDQVTLTKYGVHEIITYTMASLMNKDQIISGCGFMFEFEFLKNLRKPGDEM
EPKEEFAKFNALELDSDSLAIFLAVITLISDQRGILNVPDIETIDNLLQALELOLK
LNHSESLDFAKYLOKMTDLRQIYIYTHVOLLHVIKRTETMSLHPLQELIKDLY"

BASE COUNT      531 a      412 c      387 g      450 t
ORIGIN

Alignment Scores:
Pred. No.:      1,236-153      Length:      1780
Score:          2153.00      Matches:      424
Percent Similarity: 85.74%      Conservative: 9
Best Local Similarity: 83.96%      Mismatches: 10
Query Match:      61.37%      Indels:      62
DB:              10      Gaps:      1

US-09-931-007A-1 (1-688) x MMU09138 (1-1780)

OY      1 MetGlygluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
|||||
|||||
|||||
|||||
|||||
DB      40 ATGGCTGAACACTCTGGAGATTCTCTGTTGACCCAGACATGTCGCTTCGTCGATGCA 99
|||||
|||||
|||||
|||||
|||||
OY      21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
|||||
|||||
|||||
|||||
|||||
DB      100 CTGGCTATGAGCATTCCACAAAGAAATTCACATGCTTGACACAGACATCCATTCGGGCC 159
|||||
|||||
|||||
|||||
|||||
OY      41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
|||||
|||||
|||||
|||||
|||||
DB      160 ACCAATTGCGGATAGCTCTGTGACACTCTTCGCGATGGAAGACACTCCGATTCCTTT 219
|||||
|||||
|||||
|||||
|||||
OY      61 AspIleIysProPheThrThrValAspPheSerIleSerThrProHisTyrGluAsp 80
|||||
|||||
|||||
|||||
|||||
DB      220 GACATCAAGCCCTTACCACAGTTGATTTCTCCACACATTTCTGCTCCACACTATGAAGAC 279
|||||
|||||
|||||
|||||
|||||
OY      81 IleProIleThrThrAspProValAlaAlaAspTyrIlyTyrAspLeuIysLeuGln 100
|||||
|||||
|||||
|||||
|||||
DB      280 ATTTCATTCCACAGAGCGTCGACCAATCGCTTCGATTCGAATATATGACCTCAAGCTCAA 339
|||||
|||||
|||||
|||||
|||||
OY      101 GluTyrGlnSerAlaIleIlysValGluProIleAspProPoloTyrTyrSerGluIysThr 120
|||||
|||||
|||||
|||||
|||||
DB      340 GAATACCAAAAGTGCATCAAGTAAAGTAAACCTGCATCTCCACCTTATTATTCGAAAAGACC 399
|||||
|||||
|||||
|||||
|||||
OY      121 GlnLeuTyrAsn----- 124
|||||
|||||
|||||
|||||
|||||
DB      400 CAGCTCTACACAGAGCGCTCATGAAGAACCTTACTCTCCATGAGGCCATTCGAGAGCCGA 459
|||||
|||||
|||||
|||||
|||||
OY      124 ----- 124
|||||
|||||
|||||
|||||
|||||
DB      460 GTCTGTGGGATAAAGCATCAAGCCTTCCACTATGACATTCATGCTTGTGAAGATGCAG 519
|||||
|||||
|||||
|||||
|||||
OY      124 ----- 124
|||||
|||||
|||||
|||||
|||||
DB      520 GGTTTTTCGACAGACCATCGATTGAAGCTTATTATATGATAGGTGATCTTAACCTCC 579
|||||
|||||
|||||
|||||
|||||
OY      125 -----ArgAsnIysCysGlnTyrCysArgPheGlnIysCysLeuAla 138
|||||
|||||
|||||
|||||
|||||
DB      580 CGGATCCACAAAGAAAGTAAATTAATAGTCAGTACGTGCGGTTTCAGAACTGCGCTTGC 639
|||||
|||||
|||||
|||||
|||||
OY      139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnIleAlaGluIysGluLys 158
|||||
|||||
|||||
|||||
|||||

```

```

Db      640 GTGGGAGTCTCACAATGCCATCAGGTTTGGCGGATGCCAGGCCGAGAGGAGAG 699
Oy      159 LeuLeuAlaGluIleSerSerAspIleAspGluLeuAsnProGluSerAlaAspLeuArg 178
Db      700 CTGTTGGCGGAAATCTCCAGTATATCGACCGAGCTGAACCCAGAGTCTCTATCTGGCA 759
Oy      179 AlaLeuAlaIleHisLeuThrAspSerTyrIleLeuSerPheProLeuThrLysAlaLys 198
Db      760 GCCCTGGCAAGCATTTGATGACATCATCAATAAAGTCTCTCCGCTGCACCAAGCAAG 819
Oy      199 AlaAlaGAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
Db      820 GCGAGGGCGATCTTACAGAGAAAGCAAGGACAATCCATTGTCATCTACGACAG 879
Oy      219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
Db      880 AATTCCTTATATATGGAGAAATCAATCAAGTTCAAACATATCACCCCCTGCAGAG 939
Oy      239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
Db      940 CAGACCAAAAGAGGTGGCCATCCGAATTTTCAAGGGTGCACATTCGATCGTAGAGCC 999
Oy      259 ValGlnGluIleThrGluThrLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
Db      1000 GTGCAGAGATCACAGAGATATGCCAAAATATCCCTGGTTTCAATTAACCTGATTTGAAT 1059
Oy      279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuLysSer 298
Db      1060 GACCAAGTACCTCTGCTCAAGTATGCTGTCATGAGATCATCTACACGATGCTGGCTCC 1119
Oy      299 LeuMetAsnLysAspGlyValLeuLeuIleSerGlnGlnGlnGlyPheMetThrArgLysPhe 318
Db      1120 CTGATGAATAAAGATGAGATCCTCATCTCAGAGGGCCAAAGATCATGACAGGGAGTTC 1179
Oy      319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db      1180 CTCAAAAACCTCCGGAGAGCCCTTGTGCTTATGAGACCTTAAGTTTGAAGTTTGTGTG 1239
Oy      339 LysPheAsnAlaLeuGlnLeuLeuAspSerAspSerAspLeuAlaIlePheIleAlaValIleIle 358
Db      1240 AAGTTCAATGACACTGGAATTAATGACAGTGACTTGGCTATTTTATAGCTGTCAATTAT 1299
Oy      359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db      1300 CTCATGTGAGACCGCCAGAGGCTTGTGAACGTGAAGCCCATCGAGACATCCAAAGCAAC 1359
Oy      379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluLysSerGlnLeuPhe 398
Db      1360 CTGCTGCAGGCGCTGAACTGACGCTCAAGCTGAATCACCCAGAGTCCCTCAGCTGTTTC 1419
Oy      399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
Db      1420 GCCAAGGTGCTCCAGAGATGACAGACCTCAGAGATGTGCACAGAGACGAGCTGCAGCTA 1479
Oy      419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
Db      1480 CTGCATGTGATCAAGAGACAGAGACAGACATGAGACCTTCAACCCCTGCTCCAGAGATC 1539
Oy      439 TyrLysAspLeuTyr 443
Db      1540 TTAACAAGACTGTAT 1554

```

```

RESULT 14
AB019561      1518 bp      mRNA      linear      ROD 08-JAN-1999
LOCUS      AB019561
DEFINITION      Rattus norvegicus mRNA for PPAR gamma2, complete cds.
ACCESSION      AB019561.1 GI:4115708
VERSION      PPAR gamma2.
KEYWORDS      Rattus norvegicus (strain: Sprague-Dawley) 8-week-old male adipose
SOURCE      tissue cDNA to mRNA.
ORGANISM      Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
AUTHORS      Rattus.
TITLE      1. (sites)
JOURNAL      Tanaka,T. and Itoh,H.
REFERENCE      Down-regulation of PPAR gamma
AUTHORS      Unpublished
TITLE      2 (bases 1 to 1518)
JOURNAL      Itoh,H., Tanaka,T. and Nakao,K.
REFERENCE      Submitted
TITLE      School of Medicine, Dept of Med and Clin Sci; 54 Shogoin
JOURNAL      Kawahara-cho, Sakyo-ku, Kyoto 606-8507, Japan
              (E-mail: hito@kuhp.kyoto-u.ac.jp, Tel:81-75-751-3170,
              Fax:81-75-771-9452)
FEATURES
source      Location/Qualifiers
              1..1518
              /organism="Rattus norvegicus"
              /strain="Sprague-Dawley"
              /db_xref="taxon:10116"
              /sex="male"
              /tissue_type="adipose tissue"
              /dev_stage="8-week-old"
              1..1518
              /codon_start=1
              /product="PPAR gamma2"
              /protein_id="BAA36485.1"
              /db_xref="GI:4115709"
              /translation="MGETLGDPDPDPEHGAFADALPSTSOETIMWDTENPFWTNG
              ISSVDLSVMDHSHSFDIKPTTYDVDFSSISAPHEIDIEFTADPMVADYKDLQLEY
              QSAIKVEPASPYPVSEKTOLYNRIHERPSNSLMAIEQVCGDKSGHYGVHACEGK
              GFPRFTIRLKLIDRCDLNCRIHKSRNYCQYCFQFCLAVGSHNMIREFRMPQAK
              EKLAEISSDIDOLNPESADLRALAKHLUDYSIKSFPLTKAKARAILIKTTDKSPY
              IYDMSNLMMGEDKIKFKIIPLOBSKEVATIRFOGQCFREVAQITTEYAKRIPF
              INLDNNOYVTLIKGYHEIITYTLMASLNNKGVILSNGQCFTEFELKSLKPRGDM
              EPKFEFAVFNALDDSDLFIATVAILISDRGLNVKPTIEDIQNLQALELQLK
              LNHPSSOLFVKVLOKMTDILROIYETHVOLHLVTKRETDMSLHPLQLEYKDLV"
BASE COUNT      422 a      379 c      358 g      359 t
ORIGIN
Alignment Scores:      2,42e-153      Length:      1518
Pred. No.:      2148.00      Matches:      423
Score:      85.54%      Conservative:      9
Best Local Similarity:      83.76%      Mismatches:      11
Query Match:      61.23%      Indels:      62
DB:      10      Gaps:      1
US-09-931-007a-1 (1-688) x AB019561 (1-1518)
Oy      1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
Db      1 ATGGGGAACACTCTGGAGATCTCTCTGTGACCCAGAGCATGAGTCCCTGCCTATGCA 60
Oy      21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetCProPheThrPro 40
Db      61 CTGCCTATGAGCACTTCAAGAAATTAACCATGTGTGACACAGAGATGCATTTGCGGCC 120
Oy      41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerIleSerPhe 60
Db      121 ACCAAGCTTGGAGATCAGCTCTGTGAGCTCTCTGTGATGAGATGACACCTCCATTCCTTT 180
Oy      61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisIleGluAsp 80
Db      181 GACATCAAAACCTTTACCAACGCTGATTTCTCAGACATTTCTGCTCCACACATATGAAGAC 240
Oy      81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100
Db      241 ATCCCGTTTCAACAAGAGCTGACCAATGCTGTGATTACAAATATGACCTGAAGCTCAA 300
Oy      101 GUTYrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120
Db      301 GAATACCAAAAGTGCATCAAAAGTACAGCCTGCGTCCCGCTTATTTATTTCAAAAAACC 360

```

```

Qy 121 GlnLeuTyrAsn-----124
    |||||
Db 361 CAACTCTACAAAGGCGCACATGAAGAGCCTTCAAACTCCTCATGCGCATGAGTGGCGGA 420
Qy 124 -----124
Db 421 GTCGTGGGGATTAAGCATCAGGCTTCCACTATGAGAGTCATGCTTGTGAAGATGCAAG 480
Qy 124 -----124
Db 481 GGTTTTCCAGAGAACCATCCGATTGAAGCTTATTATGATAGGTGATCTTAAGTCT 540
Qy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
    |||||
Db 541 CGGATCCACAAAAGAGATAGAATAATGTCAGTACGTCTCGGTTTCAGAAAGGCTTCTCT 600
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
    |||||
Db 601 GTGGGGATGTCTCACAAAGCCATCAGGTTTGGGCAATGCCACAGCGCGAGAGAGAGAG 660
Qy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
    |||||
Db 661 CTGTTGGCGGAGATTCACAGTATATCACAAGCTGAACCCAGAGTCTGCTGATCTGGCA 720
Qy 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
    |||||
Db 721 GCCCTGGCAAGCATTTGTATGACTCATACATAAAGCTCTCCGTCGACCAAGCCCAAG 780
Qy 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
    |||||
Db 781 GCGAGGGGATCTTCACAGGAAGCAACACAAATACCATTTGTCATCTACACATG 840
Qy 219 AsnSerLeuMetMetGlnGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238
    |||||
Db 841 AATTCCTTAATGATGGAGGAAGACAAATCAAGTTCAACATATACCCCTCGAGGAG 900
Qy 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
    |||||
Db 901 CAGACCAAGAGGTGGCCATCCGCTTTTCAAGGCTCCAGTTTCGATCCGTGGAAGCT 960
Qy 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
    |||||
Db 961 GTGCAAGAGATCACAGAGATATGCCAAAATATCCCTGTTTCATTAACCTTGACTTGAAT 1020
Qy 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
    |||||
Db 1021 GACCAAGAGACTCTGCTCAAGTATGCTTCCATGAGATCATCTACACCATGCTGCTCC 1080
Qy 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
    |||||
Db 1081 CTGATGATTAAGAATGAGTCTCTCATATCAGAGGCAACAAAGATTGATGACAGGAGTTC 1140
Qy 319 LeuLysSerLeuAlaGlyLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
    |||||
Db 1141 CTCAAAAACCCGCGGAACCCCTTGGTACTTATGAGACCTTAAGTTGAGTTTGTCTGTG 1200
Qy 339 LysPheAsnAlaLeuGlnLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIle 358
    |||||
Db 1201 AAGTTTCATGACACTGGAATTGATGACAGTACCTTGCGCATATTTATAGCTCTATTATT 1260
Qy 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
    |||||
Db 1261 CTCATGTGAGACCGCGCGCTGCTGTAAGCTGAAGCCCATCGAGAGATCCAGAGCAAC 1320
Qy 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluLysSerGlnLeuPhe 398
    |||||
Db 1321 CTGCTGCGAGGCCCTGGAATCTCAGCTGAAGCTGAACCCCGGAGTCTCTCCACAGTGTTC 1380
Qy 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGlnHisValGlnLeu 418
    |||||
Db 1381 GCCAAGAGTCTCCAGAGATGACAGAACTCTCAGGCAAGATTGTTCACAGAGACAGTGCAGCTA 1440
Qy 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438

```

```

Db 1441 CTGCATGTGATCAAGAGAGAGAGACAGATATGAGACCTTACACCTCTGCTCAGAGATC 1500
Qy 439 TyrLysAspLeuTyr 443
    |||||
Db 1501 TACAAGGACTTGTA 1515

RESULT 15
AF156666
LOCUS
DEFINITION
Rattus norvegicus peroxisome proliferator-activated receptor gamma
2 (PPARGamma2) mRNA, complete cds.
ACCESSION
AF156666
VERSION
AF156666.1 GI:5107414
KEYWORDS
SOURCE
Rattus norvegicus.
Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 (bases 1 to 1785)
Guardiola-Diaz,H.M., Rehmark,S., Usuda,N., Albrektzen,T.,
Feltkamp,D., Gustafsson,J.A. and Alexson,S.E.
Rat peroxisome proliferator-activated receptors and brown adipose
tissue function during cold acclimatization
J. Biol. Chem. 274 (33), 23368-23377 (1999)
MEDLINE
99367468
PUBMED
10438514
REFERENCE
2 (bases 1 to 1785)
Guardiola-Diaz,H.M., Rehmark,S., Usuda,N., Albrektzen,T.,
Feltkamp,D., Gustafsson,J.A. and Alexson,S.E.H.
Direct Submission
Submitted (03-JUN-1999) Division of Clinical Chemistry, Karolinska
Institute, C1-74 Huddinge Hospital, Huddinge S-141 86, Sweden
TITLE
JOURNML
Location/Qualifiers
FEATURES
source
1..1785
/organism="Rattus norvegicus"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
1..1785
/gene="PPARGamma2"
49..1566
/gene="PPARGamma2"
/note="nuclear receptor"
/codon.start=1
/product="peroxisome proliferator-activated receptor gamma
2"
/protein_id="AAD40119.1"
/db_xref="GI:5107415"
/translation="MGETLGDPPVDPPEHGFADLPKSTSOEITMVDEMPPTNFG
ISSVDLSVMDHSHSPDIKPTTVDPSISAPHYEDIPTRADPMVADYKDLKLOEY
OSAKVPPASPPYTSKTOYLNRHPEPSNMAIERVCGDRASGTHYGVHACEGK
GFERRTIRKLITDRCDLNCRIHKSNNKQCYCFQCLAVGMSHNAIRGKMPQAK
EKLALISSDIDLPESADLRALAKLYDSYIKSPFLTAKARALITGTTDKSPV
IYDMNSLMMEGEDIKFKHITPLOSKEVAIRIFOGQSFVEVACETITEYAKNIPGF
INDLNDQVTLKKGVEIITYTMAISLMDKGVLSISGSGFMREFFLSRKPGDFM
EPKPEFAKFNALTELDDSDLAIFLAVIILISGDRGLLVNKPDIEDIONLQALTELQK
LNHPESQALFKALQKMTDLROYTEHVQLLHVTKETDMSLHPLQLEIYKDIY"
BASE COUNT
533 a 417 c 395 g 440 t
ORIGIN

Alignment Scores:
Pred. No.: 2,95e-153 Length: 1785
Score: 2148.00 Matches: 423
Percent Similarity: 85.54% Conserved: 9
Best Local Similarity: 83.76% Mismatches: 11
Query Match: 61.23% Indels: 62
DB: 10 Gaps: 1
US-09-931-007a-1 (1-688) x AF156666 (1-1785)
Qy 1 MetGlyuThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20
    |||||

```

D	b	49	ATGGGTGAAC	TCTGGGAGAT	CCCTCCGTTG	AGACCA	GAGACAT	AGTGCCTT	CGCTATG	CA	108		
O	y	21	LeuSerAla	asnIle	SerGln	gluMet	ThrMet	ValAsp	ThrGlu	MetPro	ThrPro	40	
D	b	109	CTGGCTAT	GACACTT	CAACAAGAA	TAATAC	CATGGTTG	ACACAG	AGATGC	CAATTC	TGGGCC	168	
O	y	41	Thr	asnPhe	GlyIle	SerSer	ValAsp	LeuSer	ValMet	Glu	aspHis	SerHis	60
D	b	169	ACCACTT	GGAATC	AGCTCTG	TGGACCT	CTCTG	GTATGG	ATACAC	CACTCC	CAATTC	228	
O	y	61	Asp	Ile	Val	ProPhe	ThrVal	AspPhe	SerSer	Ile	SerThr	ProHis	80
D	b	229	GACAT	CAACAC	CCCTTACC	ACGGT	GATTTCT	CCAGCAT	TTTCTCT	CCACAC	ATATGAC	288	
O	y	81	Ile	Pro	Phe	Thr	Arg	Thr	Asp	Pro	ValAla	Asp	100
D	b	289	ATCCG	GTTCAC	AAAGAG	CTCCAC	CAATGG	TGTGC	ATTAAC	AAATAT	GACCTG	348	
O	y	101	Glu	Thr	Gln	Ser	AlaIle	ValGlu	Pro	AlaSer	Pro	Thr	120
D	b	349	GAA	ATCCAA	AGTGC	ATCAAG	TAGAC	CCCTG	CCGCCCT	CTATTAT	TCTG	408	
O	y	121	Gln	Ileu	Thr	Asn						124	
D	b	409	CAACT	CTAC	CAACAG	CCAC	CAATG	GAAG	ACCTTCA	AACTCC	CTCAT	468	
O	y	124	-----	-----	-----	-----	-----	-----	-----	-----	-----	124	
D	b	469	GTC	TGTGG	GTAA	AGCAT	CAGCTT	CCACT	ATGAG	AGTCC	ATGTTGA	528	
O	y	124	-----	-----	-----	-----	-----	-----	-----	-----	-----	124	
D	b	529	GGT	TTTTT	CCG	AGAAC	CAATCC	ATGAA	CTTATT	ATGAT	AGGTG	588	
O	y	125	-----	-----	-----	-----	-----	-----	-----	-----	-----	138	
D	b	589	CGAT	CCCA	AAAGAG	TAGAA	TAATAT	GTCA	CTACTG	CTTC	CAGAA	648	
O	y	139	Val	Gly	Met	Ser	His	Asn	AlaIle	Arg	Phe	Gly	158
D	b	649	GTC	GGGAT	GTCTC	ACATG	CCATC	AGTGG	CGCAAT	CCCA	GAGCG	708	
O	y	159	Leu	Ileu	Ala	Glu	Ile	Ser	Ser	Asp	Ile	Asp	178
D	b	709	CTG	TGGG	GGGAG	ATCTC	CAATG	ATATG	ACAC	CACTAA	CCAGAG	768	
O	y	179	Ala	Leu	Ala	Val	His	Ser	Thr	Asp	Ser	Thr	198
D	b	769	GCC	CTG	GAAG	CACTT	TGTAT	GACTC	ATAC	ATAAT	GCTTCC	828	
O	y	199	Ala	Arg	Ala	Ile	Leu	Thr	Gly	Val	Thr	Asp	218
D	b	829	GCG	AGG	GGCAT	CTTG	ACAG	GAAG	ACAC	AGCA	CAAT	CACT	888
O	y	219	Asn	Ser	Leu	Met	Gly	Glu	Asp	Val	Ile	Gly	238
D	b	889	AAT	TCCTT	AA	GATG	GGAG	GAAG	CAAA	TAAT	CAAT	CA	948
O	y	239	Gln	Ser	Val	Glu	Ala	Ile	Arg	Gly	Phe	Gln	258
D	b	949	CAG	AGCA	AAAG	GGGCC	ATCCG	CAATTTT	CAAA	GGGTG	CCACTT	1008	
O	y	259	Val	Gln	Glu	Ile	Thr	Gly	Val	Ala	Val	Ser	278
D	b	1009	GTC	CAAG	ATC	ACAG	ATGTA	TGCA	AAATAT	ATCC	GTTC	1068	
O	y	279	Asp	Gln	Val	Thr	Leu	Leu	Val	Thr	Gly	Val	298
D	b	1069	GAC	CAAG	ATC	ATCTG	CTCA	AGTAT	GTG	TCC	ATGAT	1128	
O	y	299	Leu	Met	Val	Ser	Asp	Gly	Val	Leu	Ile	Ser	318
D	b	1129	CTG	ATG	ATTA	AGTAT	GAGTCT	CTCAT	TAC	AGAG	CAAG	1188	

QY	319	LeuysSerLeuAtgIysrProPheGlyAspPheMetGluProIlysPheGluPheLeuAla	338
Db	1189	CTCAAAAGCCCTGGCGAAGCCCTTTGGTGGATTTATGTGACCCTAAATTTGAGTTGCTGTG	1248
QY	339	LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle	358
Db	1249	AAGTTTCANTGCACCTGGAAATTAGATGACAGTGAAGCTTGGCCATTATTTATAGCTGTCTATT	1308
QY	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGluAspAsn	378
Db	1309	CTCAGTGGAGACCGCCAGGCTGTGTGAACGTGAACCCATCGAGACATCCAAAGACAAAC	1368
QY	379	LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe	398
Db	1369	CTGTGTGCAGGGCCCTGGAACTCCAGCTGAAGCTGAACCCAGGATCTCTCCAGCTGTTC	1428
QY	399	AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu	418
Db	1429	GCCAGGTGCTCCAGAAAGATGACAGACCTTCAGCAATTTGTACACAGAGCACCTGCAGCTTA	1488
QY	419	LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGlnIle	438
Db	1489	CTGCATGTGATCAAGAGAGAGGAGACAGATATGAGCTTCACACCTGTGCTCCAGAGATC	1548
QY	439	TyrLysAspLeuTyr	443
Db	1549	TACAAAGCACTTGTA	1563

Search completed: February 25, 2003, 05:32:08  
Job time : 4897 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:23:46 ; Search time 75 Seconds  
(without alignments)  
1890.139 Million cell updates/sec

Title: US-09-931-007a-1  
Perfect score: 3508  
Sequence: 1 MGETLGSDSPIDPESDSFTDT.....KTEFMSLHPLQEIYKDLV 688

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues  
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: SPREMBL\_21:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteria:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2212	63.1	505	6	095K28
2	2089	59.5	477	4	096J12
3	2064	58.8	475	6	095J17
4	2045	58.3	475	11	099PC7
5	1920	54.7	475	13	091878
6	1916	54.6	472	13	080GCO
7	1698	48.4	477	13	098SE8
8	1564	44.6	393	13	099U74
9	1549	44.2	393	13	099U26
10	1325	37.8	333	13	099W01
11	1312.5	37.4	532	13	09W712
12	1140	32.5	543	13	090WP6
13	1117.5	31.9	443	13	0918W3
14	1105	31.5	219	6	095K74
15	1089.5	31.1	468	6	095N78
16	1085.5	30.9	440	11	099ND3

17	1084.5	30.9	441	6	09M2G4	09m2g4 sus scrofa
18	1082	30.8	216	6	08WMS4	08wm54 canis famill
19	1080.5	30.8	440	11	090M10	09gw10 rattus norv
20	1078.5	30.7	521	13	0918F6	0918f6 salmo salar
21	1076.5	30.7	368	13	090266	090266 brachydanio
22	1076.5	30.7	440	11	062879	062879 rattus norv
23	1076.5	30.7	441	6	09N2H4	09n2h4 oryctolagus
24	1070.5	30.5	337	13	080UX1	08ux1 oryzias lat
25	1070	30.5	468	13	080GCL	08ggcl anser anser
26	1068.5	30.5	468	13	0918W4	0918w4 gallus gall
27	1067.5	30.4	436	6	09N128	09n128 bos taurus
28	1065.5	30.4	470	6	09N129	09n129 bos taurus
29	1056	30.1	328	13	09P076	09p076 crocodylus
30	1035.5	29.5	329	13	090267	090267 brachydanio
31	1033	29.4	467	11	070527	070527 cavia porce
32	1028.5	29.3	468	6	09N135	09n135 sus scrofa
33	1021.5	29.1	337	13	090262	090262 salmo salar
34	1020.5	29.1	304	13	09P075	09p075 crocodylus
35	1000	28.5	252	6	046479	046479 oryctolagus
36	969	27.6	482	13	0918F7	0918f7 salmo salar
37	935	26.7	436	13	090Y02	090y02 petromyzon
38	802.5	22.9	361	4	09BUD4	09bud4 homo sapien
39	738	21.0	230	6	09GLC8	09glc8 mustela vis
40	715	20.4	196	13	09W718	09w718 gallus gall
41	684.5	19.5	233	13	080UX2	08ux2 pimephales
42	571.5	16.3	174	6	046477	046477 oryctolagus
43	509.5	14.5	157	4	075780	075780 homo sapien
44	394.5	11.2	549	13	090970	090970 gallus gall
45	392.5	11.2	615	11	0922A5	0922a5 mus musculu

## ALIGNMENTS

### RESULT 1

095K28 PRELIMINARY: PRT: 505 AA.  
ID 095K28;  
AC 095K28;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Peroxisome proliferator-activated receptor gamma 2.  
GN PPARGAMA2.  
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
OC Cercopithecoidea; Macaca.  
OX NCBI\_TaxID=9541;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Zhou J., Wilson K.M., Medh J.D.;  
RT "Identification of novel peroxisome proliferator-activated receptor-  
RT gamma splice variants and induction of PPAR-gamma expression by a  
RT high-fat diet in monkey macrophages."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL: AY048695; AA05262.1; -  
DR InterPro: IPR000536; Hormone\_rec\_119.  
DR InterPro: IPR001628; ZnF\_C4steroid.  
DR Pfam: PF00104; hormone\_rec\_1.  
DR Pfam: PF00105; zf-C4; 1.  
DR PRODOM: PD000035; ZnF\_C4steroid; 1.  
DR PROSITE: PS00031; NUCLEAR RECEPTOR; UNKNOWN\_1.  
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
SQ SEQUENCE 505 AA; 57624 MW; ACE6A50AE8AF694A CRC64;

Query Match 63.1%; Score 2212; DB 6; Length 505;  
Best Local Similarity 86.7%; Pred. No. 6.8e-131;  
Matches 436; Conservative 3; Mismatches 2; Indels 62; Gaps 1;  
QY 1 MGETLGSDSPIDPESDSFTDTLSANISQEMTWVDEMPFWPTNFGISSVDLSMEDHSHSF 60

```

|||||
Db 1 MGEETGDSFIDESDSFDTLSANISOETIWDTEPMPPTNFGISSVDLSVMDHSHSF 60
Oy 61 DKKPPTVDFESSITPHYEDIPFTRTDPVADYKXDKLOEYOSAIKVPASPPYSEKT 120
Db 61 DKKPPTVDFESSISAPHEDIPFTRTDPVADYKTDLOEYOSAIKVPASPPYSEKT 120
Oy 121 QLYN-----RNCQYCRFOKCLAVGSMHNAIRFGMPQAEKKLAEISSDIDQLNPSADLR 124
Db 121 QLYNPHREPSNSLMAIECRVCGDKASGFHGVHACEGCKGFRRRTIRLKLTYDRCDLNC 180
Oy 125 -----RNCQYCRFOKCLAVGSMHNAIRFGMPQAEKKLAEISSDIDQLNPSADLR 178
Db 181 RHKRSRNMCOQCRFOKCLAVGSMHNAIRFGMPQAEKKLAEISSDIDQLNPSADLR 240
Oy 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMGDEKIKFKHITPLOE 238
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPVIYDMNSLMGDEKIKFKHITPLOE 300
Oy 239 OSKEVAIRIFOGCQFRSVEAVQETREYAKSIRPGFVNLNDQVTLTKYGVHEIITYTMLS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVQETREYAKSIRPGFVNLNDQVTLTKYGVHEIITYTMLS 360
Oy 299 LMNKGCVLISEGQGFMTREFLSLKRPGDEMPKFEPAVKNNALBLDSDLAIFLAVII 358
Db 361 LMNKGCVLISEGQGFMTREFLSLKRPGDEMPKFEPAVKNNALBLDSDLAIFLAVII 420
Oy 359 LSGDRPGLLNVPRIEDIDONLQALELOELKLNHPRESSOLFALOKMTDLROIVTEHVOL 418
Db 421 LSGDRPGLLNVPRIEDIDONLQALELOELKLNHPRESSOLFALOKMTDLROIVTEHVOL 480
Oy 419 LOVIKKTETDMSLHPLLOEYKDLX 443
Db 481 LOVIKKTETDMSLHPLLOEYKDLX 505

```

## RESULT 2

```

Oy 096J12 PRELIMINARY: PRT: 477 AA.
AC 096J12:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Unknown (protein for MGC:5041).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; BC006811; AA06811.1; -.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 477 AA: 54681 MW: 1061C2074B739E0A CXC64:

```

Query Match 59.5%; Score 2089; DB 4; Length 477;  
 Best Local Similarity 87.0%; Pred. No. 3.2e-125;  
 Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

Oy 29 MMTVTEPMPPTNFGISSVDLSVMDHSHSFIDKPTTYVDFESSISPHYEDIPFTTRDP 88
Db 1 MMTVTEPMPPTNFGISSVDLSVMDHSHSFIDKPTTYVDFESSISPHYEDIPFTTRDP 60

```

```

Oy 89 VVADYKYDKLOEYOSAIKVPASPPYSEKTOLYN----- 124
Db 61 VVADYKYDKLOEYOSAIKVPASPPYSEKTOLYNPHREPSNSLMAIECRVCGDKASG 120
Oy 125 -----RNCQYCRFOKCLAVGSMHNAIRFGMPQAEKKLAEISSDIDQLNPSADLR 146
Db 121 FHYGVHACEGCKGFRRRTIRLKLTYDRCDLNCIRHKSRSNKCQYCRFOKCLAVGSMHNAIR 180
Oy 147 RFGMRPOAEKELAEISSDIDQLNPSADLRALAKHLVDSYIKSFPLTKAKARAILTGK 206
Db 181 RFGMRPOAEKELAEISSDIDQLNPSADLRALAKHLVDSYIKSFPLTKAKARAILTGK 240
Oy 207 TTDKSPVIYDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCQFRSVEAVQETREYA 266
Db 241 TTDKSPVIYDMNSLMGDEKIKFKHITPLOSKEVAIRIFOGCQFRSVEAVQETREYA 300
Oy 267 KSIQGFVNLNDQVTLTKYGVHEIITYTMLSAMKDGVLISEGQGFMTREFLSLRPF 326
Db 301 KSIQGFVNLNDQVTLTKYGVHEIITYTMLSAMKDGVLISEGQGFMTREFLSLRPF 360
Oy 327 GDFMPKFEPAVKNNALBLDSDLAIFLAVIILSGDRPGLLNVPRIEDIDONLQALELO 386
Db 361 GDFMPKFEPAVKNNALBLDSDLAIFLAVIILSGDRPGLLNVPRIEDIDONLQALELO 420
Oy 387 LKLNHPRESSOLFALOKMTDLROIVTEHVOLQYIKKTETDMSLHPLLOEYKDLX 443
Db 421 LKLNHPRESSOLFALOKMTDLROIVTEHVOLQYIKKTETDMSLHPLLOEYKDLX 477

```

## RESULT 3

```

Oy 095J17 PRELIMINARY: PRT: 475 AA.
AC 095J17:
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor gamma 3 (peroxisome
DE proliferator-activated receptor gamma 1).
GN PPARGAMMA3 OR PPARGAMMA1.
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Macaca.
OX NCBI_TaxID=9541;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhou J., Wilson K.M., Medh J.D.;
RT "Identification of novel peroxisome proliferator-activated receptor-
RT gamma splice variants and induction of PPAR-gamma expression by a
RT high-fat diet in monkey macrophages."
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL; AY048696; AA05263.1; -.
DR EMBL; AY048694; AA05263.1; -.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec.1.
DR Pfam; PF00105; zf-C4; 1.
DR ProDom; PD000035; Znf_C4steroid; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; UNKNOWN_1.
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
SQ SEQUENCE 475 AA: 54470 MW: D16C3A931EB889181 CXC64:

```

Query Match 58.8%; Score 2064; DB 6; Length 475;  
 Best Local Similarity 86.1%; Pred. No. 1.2e-121;  
 Matches 409; Conservative 2; Mismatches 2; Indels 62; Gaps 1;

```

Oy 31 MMTVTEPMPPTNFGISSVDLSVMDHSHSFIDKPTTYVDFESSISPHYEDIPFTTRDPV 90
Db 1 MMTVTEPMPPTNFGISSVDLSVMDHSHSFIDKPTTYVDFESSISPHYEDIPFTTRDPV 60

```

```

QY 91 ADYKDKLQEOYSAIKVEPASPYYSEKTLQYN-----124
D6 61 ADYKDKLQEOYSAIKVEPASPYYSEKTLQYNKPHEPEPSNSLMAIBCRVCGDKASGFH 120
QY 125 -----RNNCOYCRFOKCLAVGMSHNAIRF 148
D6 121 YGVNAECGCKGFRRRTIRLKLHYRCDLNCRIHKKSNNKCOYCRFOKCLAVGMSHNAIRF 180
QY 149 GMPQAEKEXLLAEISSDIDLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 208
D6 181 GMPQAEKEXLLAEISSDIDLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 240
QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 268
D6 241 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 300
QY 269 IPGFVNLNDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLKSIRKPPGD 328
D6 301 IPGFVNLNDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLKSIRKPPGD 360
QY 329 FMEPKFEFAVFNALIEDSDLAIFAVIILISGDRPGLLNKPRIEDIODNLQALELQLK 388
D6 361 FMEPKFEFAVFNALIEDSDLAIFAVIILISGDRPGLLNKPRIEDIODNLQALELQLK 420
QY 389 LNHPESSOLFVKLQKMTDLRQIYTEHVQLQVIKKTTDMSLHPLQEIYKDLX 443
D6 421 LNHPESSOLFVKLQKMTDLRQIYTEHVQLQVIKKTTDMSLHPLQEIYKDLX 475

```

RESULT 4  
099PC7 PRELIMINARY; PRT: 475 AA.

```

ID 099PC7 AC 099PC7:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator activated receptor-gamma (Fragment).
OS Cavia porcellus (Guinea pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Hystriognathii; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ADIPOSE TISSUE;
RA Khoo B.Y., Najimudin N., Tengku Muhammad T.S.;
RT "Molecular Cloning and Sequencing of Peroxisome Proliferator Activated
  Receptor-gamma (PPARGamma) from Guinea Pig.";
RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AF317514; AAG60685.2; -.
DR HSSP: P37231.1EM9.
DR InterPro: IPR000536; Hormone_rec_1ig.
DR InterPro: IPR001723; Stdhnm_receptor.
DR InterPro: IPR001628; Znfc4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR PRINTS: PR00398; STPDHORMER.
DR PRINTS: PR00047; STRODPTNGER.
DR PRODom: PD000035; Znfc4steroid.1.
DR SMART: SM00430; HOL1.1.
DR SMART: SM00399; znfc4.1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR.1.
DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;
KW Zinc-finger.
FT NON_TER 1
SQ SEQUENCE 475 AA; 54425 MW; DE45CF49C896F401 CRC64;

```

Query Match 58.3% Score 2045; DB 11; Length 475;  
Best Local Similarity 85.1% Pred. No. 1.9e-120;  
Matches 404; Conservative 5; Mismatches 4; Indels 62; Gaps 1;

```

QY 31 MVDTEMFPWPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISTPHYEDIPTRTDPV 90
D6 1 MVDTEMFPWPNFGISSVDLSVMDHSHSFDIKPTTVDFSSISAPRHEDIPTRADPMV 60
QY 91 ADYKDKLQEOYSAIKVEPASPYYSEKTLQYN-----124
D6 61 ADYKDKLQEOYSAIKVEPASPYYSEKTLQYNKPHEPEPSNSLMAIBCRVCGDKASGFH 120
QY 125 -----RNNCOYCRFOKCLAVGMSHNAIRF 148
D6 121 YGVNAECGCKGFRRRTIRLKLHYRCDLNCRIHKKSNNKCOYCRFOKCLAVGMSHNAIRF 180
QY 149 GMPQAEKEXLLAEISSDIDLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 208
D6 181 GMPQAEKEXLLAEISSDIDLNPESADLRALAKHLVSYIKSPFLTKAKARAILTGKTT 240
QY 209 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKS 268
D6 241 DKSPFVIYDMNSLMNGEDKIKFKHTTLPLOESKEVAIRIFOGCQFRSVEAVOEITEYAKN 300
QY 269 IPGFVNLNDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLKSIRKPPGD 328
D6 301 IPGFVNLNDLNDQVTLTKYGVHEIITYTMLASLMNKDGVLSISGQGFMTREFLKSIRKPPGD 360
QY 329 FMEPKFEFAVFNALIEDSDLAIFAVIILISGDRPGLLNKPRIEDIODNLQALELQLK 388
D6 361 FMEPKFEFAVFNALIEDSDLAIFAVIILISGDRPGLLNKPRIEDIODNLQALELQLK 420
QY 389 LNHPESSOLFVKLQKMTDLRQIYTEHVQLQVIKKTTDMSLHPLQEIYKDLX 443
D6 421 LNHPESSOLFVKLQKMTDLRQIYTEHVQLQVIKKTTDMSLHPLQEIYKDLX 475

```

RESULT 5

```

ID 091878 AC 091878:
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Peroxisome proliferator-activated receptor (Peroxisome
  proliferator-activated receptor gamma).
DE PPAR GAMMA.
GN Gallus gallus (Chicken).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato K., Akiba Y.;
RT "Characterization of a cDNA sequence encoding the peroxisome
  proliferator activated receptor gamma in Brouler chicken.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
CC [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=20267232; PubMed=10809235;
RA Takada I., Yu R.T., Xu H.E., Lambert M.H., Montana V.G., Klierer S.A.,
  Evans R.M., Umesono K.;
RA "Alteration of a single amino acid in peroxisome proliferator-
  activated receptor-alpha (PPAR alpha) generates a PPAR delta
  phenotype.";
RL Mol. Endocrinol. 14:733-740(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RA Meng H., Li J., Zhao J., Wang Q., Li H.;
RT "The SNPs detection on exon 2 of PPAR gamma gene in chicken.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.
DR EMBL: AB045597; BAA98100.1; -.
DR EMBL: AF163811; AAF80170.1; -.

```





DR Pfam; PF00105; zf-C4; 1.  
 DR PRINTS; PR000398; STRDHOMONER.  
 DR PRINTS; PR000447; STROIDFINGER.  
 DR ProDom; PD000035; ZnF\_C4steroid; 1.  
 DR SMART; SM00430; HOL1; 1.  
 DR SMART; SM00399; ZnF\_C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
 zinc-finger.  
 KW NON\_TER 1 1  
 FT NON\_TER 393 393  
 SO SEQUENCE 393 AA; 44812 MW; 115E7D61AAB19E6F CRC64;

Query Match 44.2%; Score 1549; DB 13; Length 393;  
 Best Local Similarity 79.6%; Pred. No. 1.8e-89;  
 Matches 313; Conservative 9; Mismatches 9; Indels 62; Gaps 1;

OY 105 AIKVPASPPYSEKTLQYN-----RNKCYCRFOKCLAVGSMNAIRFGMRPOAEKELAE 124  
 DB 1 AIKMEPPSPYFSEKVLQYLNKRPETNSLMAIEGRVCGDKASGFHYGVHACEGCKGFEFR 60  
 OY 125 -----RNKCYCRFOKCLAVGSMNAIRFGMRPOAEKELAE 162  
 DB 61 RTIRKLIVRCDLNCRIHKSRNKCQYCRFOKCLAVGSMNAIRFGMRPOAEKELAE 120  
 OY 163 ISSDIDQNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMSLM 222  
 DB 121 ISSDIDQNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMSLM 180  
 OY 223 MGEDEIKRKHITRLODOSKEVAIRIFGCGQFRSVAVOETITAYASIGFVNLNDQYT 282  
 DB 181 MGEDEINRKHVTRLODOSKEVAIRIFGCGQFRSVAVOETITAYASIGFVNLNDQYT 240  
 OY 283 LKQVYHEITVTLASLANKDGVLLISEGQFMTREFLSLRKPFDFMPPREFAVKENA 342  
 DB 241 LKQVYHEITVTLASLANKDGVLLISEGQFMTREFLSLRKPFDFMPPREFAVKENA 300  
 OY 343 IELDDSDLAIFLAVIILISGDRPGLLVKPIEDIQNLLQALELQKLNHPRESSQLFAKL 402  
 DB 301 IELDDSDLAIFLAVIILISGDRPGLLVKPIEDIQNLLQALELQKLNHPRESSQLFAKL 360  
 OY 403 QKMTDLRQIVTEHVQLLQYIKKTEIDMSLHPL 435  
 DB 361 QKMTDLRQIVTEHVQLLQYIKKTEIDMSLHPL 393

RESULT 10  
 O9PW01 PRELIMINARY: PRT; 533 AA.

AC O9PW01; 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN Peroxisome proliferator-activated receptor gamma.  
 OS Pleuronectes platea (Plaice).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 OC Pleuronectoidei; Pleuronectidae; Pleuronectes.  
 OX NCBI\_TaxID=8262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Gill;  
 RA Leaver M.J.;  
 RT "A peroxisomal proliferator-activated receptor gamma gene from a  
 marine fish";  
 RL Submitted (Aug-1996) to the EMBL/GenBank/DBJ databases.  
 CC -I- SUBCELLULAR LOCATION: NUCLEUS (BY SIMILARITY).  
 CC -I- SIMILARITY BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL; AJ243956; CAB51618.1; -.  
 DR HSSP; P37231; 1FW9.  
 DR InterPro; IPR000536; Hormone\_rec\_1lg.

DR InterPro; IPR001723; Stdhmn\_receptor.  
 DR InterPro; IPR001628; ZnF\_C4steroid.  
 DR Pfam; PF00104; hormone\_rec; 1.  
 DR Pfam; PF00105; zf-C4; 1.  
 DR PRINTS; PR00398; STRDHOMONER.  
 DR PRINTS; PR000447; STROIDFINGER.  
 DR ProDom; PD000035; ZnF\_C4steroid; 1.  
 DR SMART; SM00430; HOL1; 1.  
 DR SMART; SM00399; ZnF\_C4; 1.  
 DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
 zinc-finger.  
 KW NON\_TER 1 1  
 FT NON\_TER 533 533  
 SO SEQUENCE 533 AA; 60513 MW; DD3955E777961300 CRC64;

Query Match 37.8%; Score 1325; DB 13; Length 533;  
 Best Local Similarity 50.6%; Pred. No. 3e-75;  
 Matches 276; Conservative 59; Mismatches 66; Indels 144; Gaps 6;

OY 31 NYDT-EMPPWPTNFGISSVDLSVMEDSHSPDIKFTTVDFSSIST----- 75  
 DB 1 NYDTQQLFWPVGFSLASVDLSLDDSSHSLDKHILATIDYTSISASVPSLSFQLMSS 60  
 OY 76 -----PHYEDIPF-----RTDPVADYKTKLQEQYSAIKVPPASP 114  
 DB 61 ISSVGMAVDPSPPOSEEHLTNDYTNMHSYRTEPV-----HNSIKMEPESP 108  
 OY 115 YYSSE----- 118  
 DB 109 QYSDSPVFSKLDDDTTASLNIETECVCGDKASGFHYGVHACEGCKGFFRRTRILKLVYDH 168  
 OY 119 -----KTQLYNKNKCQYCRFOKCLAVGSMNAIRFGMRPOAEKELAEISSDIDQNP 173  
 DB 169 CDLHCRIRHKSRNKCQYCRFOKCLAVGSMNAIRFGMRPOAEKELAEFSSDMCHMPE 228  
 OY 174 SADLRALKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMSLMGEDIKFKHI 233  
 DB 228 AADLRALKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFYIYDMSLMGEDIKFKHI 288  
 OY 234 TPLODSK-----EVAIRIFGCGQFRSVA 258  
 DB 269 MPIDGQQAQSVLTATNRGLTEHHMGSDYGVWGTSTISGQEPONALELRFPOCQSRSAEA 348  
 OY 318 VOETIEYAKSITGFVNLNDQYTLTKGYMEITITMLASLANKGVLLISEGQFMTREF 318  
 DB 349 VREYTEFAKSIPTFDLNDQYTLTKGYVLEVLITMMSPLMNKGTSTISGQFMTREF 408  
 OY 319 LKSLRPPGDFMEKPEFAVKFNALDSDLAIFLAVIILISGDRPGLLVKPIEDIQN 378  
 DB 409 LKSLRPPGDFMEKPEFAVKFNALDSDLAIFLAVIILISGDRPGLLVKPIEDIQN 468  
 OY 379 LQALELQKLNHPRESSQLFAKLQKMTDLRQIVTEHVQLLQYIKKTEIDMSLHPLQEI 438  
 DB 469 VLHSELQKLNHPRESSQLFAKLQKMTDLRQIVTEHVQLLQYIKKTEIDMSLHPLQEI 528  
 OY 439 YKDL 443  
 DB 529 MKDLY 533

RESULT 11  
 O9W712 PRELIMINARY: PRT; 532 AA.

AC O9W712; 01-NOV-1999 (TREMBLrel. 12, Created)  
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
 DE 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 GN Peroxisome proliferator-activated receptor gamma.  
 OS Pleuronectes platea (Plaice).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
 OC Pleuronectoidei; Pleuronectidae; Pleuronectes.

OX NCBI\_TaxID=8262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Leaver M.J.;  
 RT "A peroxisome proliferator-activated receptor gamma gene from a marine fish."  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL: AJ249075; CAB51396.1; -  
 DR EMBL: AJ249076; CAB51396.1; JOINED.  
 DR EMBL: AJ249077; CAB51396.1; JOINED.  
 DR EMBL: AJ249078; CAB51396.1; JOINED.  
 DR EMBL: AJ249079; CAB51396.1; JOINED.  
 DR EMBL: AJ249080; CAB51396.1; JOINED.  
 DR HSSP: P37231; IEM9.  
 DR InterPro: IPR000536; Hormone\_rec\_1ig.  
 DR InterPro: IPR001628; ZnF\_C4steroid.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; zf-C4; 1.  
 DR PRINTS: PR00047; STROIDFINGER.  
 DR ProDom: PD000035; znf\_C4steroid; 1.  
 DR SMART: SM00430; HOLT; 1.  
 DR SMART: SM00399; ZnF\_C4; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; 1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.  
 SQ SEQUENCE 532 AA; 60176 MW; A3E216B79BDC7E9D CRC64;

Query Match 37.4%; Score 1312.5; DB 13; Length 532;  
 Best Local Similarity 50.5%; Pred. No. 1.8e-74;  
 Matches 275; Conservative 60; Mismatches 65; Indels 145; Gaps 7;

QY 31 MVDT-EMFVFPNGISSVDLSVMDHSHSPDIKFTVDDSSIST----- 75  
 Db 1 VMDTQQLLAWPVGSLSAVDSLEDDSSHSIDMKRLATLDYTSISSASVPSLSPOLMSS 60  
 QY 76 -----PRYEDIPFT-----RTDPVADYKYDLKLOEYQSAIKVEPASP 114  
 Db 61 ISPVGMAYDPSPOSEEHLTMDYTNMHSYTERPV-----HNSIKMEPESP 108  
 QY 115 YYSF----- 118  
 Db 109 QYSDSPVFSKLDDPTAASLNIECRVCGDKASGFHYGVHACEGCKGFPRRTIRLKYVDH 168  
 QY 119 -----KTOLYRNKCOYCRFOKCLAVGSHNAIRGMRPQAEKEKLLAESDIDLNP 173  
 Db 169 CDLHCRHKSRNKCQYCRFOKCLAVGSHNAIRGMRPQAEKEKLLAESDMEHMRPE 228  
 QY 174 SADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTDKSPFIYDMSLMGEDKIKFKHI 233  
 Db 229 AADLRALAKHLYENLYKFPPLTKAKARAILSGKTGDNAFPIYHDKMSLMEEQFNTCMQ 288  
 QY 234 TPLOEOST-----EVAIRIFOGCOFRSVEA 258  
 Db 289 -PIEOEQASVLTAAHGSLTEVHMGSIDYGVGMTSISGEPONALELRFQSCQSPSAEA 347  
 QY 259 VOETTEVAKSIPGFVNLDLNDQVTLTKYGVHEIITYTMASLMNKGVLISGQGFMTREF 318  
 Db 348 VREYTEFETKSIPTGFTDLNDQVTLTKYGVLEVLIMSPMLNKGVLISGQGFMTREF 407  
 QY 319 LKSLRKPGDMEPRFEFAVAFKNALELDDSDLAIFIAVITILSGRPGILANKPREDIDN 378  
 Db 408 LKSLRKPGDMEPRFEFAVAFKNALELDDSDMALEFVITILSGRPGILANKPREDIDN 467  
 QY 379 LLQALELQKLNHPSSOLFAPKLLQKMTDLRQIVTEHVOLLQVTKTETMSLHPLLOEI 438  
 Db 468 VLHSLLELQKLNHPSSOLFAPKLLQKMTDLRQIVTEHVOLLQKLTETMSLHPLLOEI 527  
 QY 439 YKDLT 443  
 Db 528 MKDLT 532

RESULT 12  
 Q90WP6 PRELIMINARY; PRT; 543 AA.  
 AC Q90WP6  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
 DE Peroxisomal proliferator-activated receptor gamma.  
 GN PPAR GAMMA.  
 OS Salmo salar (Atlantic salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.  
 OX NCBI\_TaxID=8030;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Fontagne S., Ezaz T., Tocher D.R., Leaver M.J.;  
 RT "Isolation and characterisation of Atlantic salmon PPAR genes."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
 DR EMBL: AJ416951; CAC95230.1; -  
 DR InterPro: IPR000536; Hormone\_rec\_1ig.  
 DR InterPro: IPR001628; ZnF\_C4steroid.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; zf-C4; 1.  
 DR ProDom: PD000035; znf\_C4steroid; 1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR; UNKNOWN\_1.  
 DR DNA-binding: Nuclear protein; Receptor; Transcription regulation;  
 KW Zinc-finger.  
 SQ SEQUENCE 543 AA; 62509 MW; 2559EC2497P58147 CRC64;

Query Match 32.5%; Score 1140; DB 13; Length 543;  
 Best Local Similarity 45.1%; Pred. No. 1.2e-63;  
 Matches 246; Conservative 65; Mismatches 100; Indels 134; Gaps 7;

QY 31 MVDTEMPFP-TNFGISSVDLSVMDHSHSPDIKFTVDDSSIST-----SITPFIYDIP 82  
 Db 1 VMDTRRAWSLISGLGLTLDVEMDKNSFDMKTLSTLDIPIYPLSLYSNNSHHHNSP 60  
 QY 83 -----FTTDPVADYKYDL 97  
 Db 61 DRSHSCNHSPPDRSHSFNHSPPDRSHSFNHSPPDRSHSFNHSPPDRSHSFNHSPPDRSH 118  
 QY 98 KLOEYQSAIKVEPAS-----PRYSKTOLYN----- 124  
 Db 119 TYSVYQGSVNDKPLSPSQSDCSIVSLSPRPHSNPPTYTDASSLNLDRCVCGDKASGF 178  
 QY 125 -----RNKCQYCRFOKCLAVGSHNAIR 147  
 Db 179 HYGVHVEGCGCGFFRRYRLKLYVDHCDLHCRHKSRNKCQYCRFOKCLAVGSHNAIR 238  
 QY 148 FGMRPQAEKEKLLAESDIDLNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKT 207  
 Db 239 FGMRPQAEKEKLLAEFMDVPRNPESADLRALRQCLSYHNRHPLTKSAKAILSGKT 298  
 QY 208 TDKSPFIYDMSLMGEDK-----KFKHITPLOEOSTKVAIRIFOGCOFRSVEA 258  
 Db 299 HGNSPFIYHDKSLTAGQYFNCROLRYLERORSVLPPEEPAEELISVFRIOFRSAEA 358  
 QY 259 VOETTEVAKSIPGFVNLDLNDQVTLTKYGVHEIITYTMASLMNKGVLISGQGFMTREF 318  
 Db 359 VOETTEVAKSIPGFTELDMDNQVTLTKYGVLEVTMTAPLPMNKGVLISGQGFMTREF 418  
 QY 319 LKSLRKPGDMEPRFEFAVAFKNALELDDSDLAIFIAVITILSGRPGILANKPREDIDN 378  
 Db 419 LKSLRKPGDMEPRFEFAVAFKNALELDDSDMALEFVITILSGRPGILANKPREDIDN 478  
 QY 379 LLQALELQKLNHPSSOLFAPKLLQKMTDLRQIVTEHVOLLQVTKTETMSLHPLLOEI 438  
 Db 479 VLQALELQKLNHPDQOLFAPKLLQKMTDLRQIVTEHVOLLQKLTETMSLHPLLOEI 538



OY 439 YKDL 443  
DB 539 MRDL 543

## RESULT 13

OY18W3 PRELIMINARY: PRT: 443 AA.  
AC Q918W3;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Peroxisome proliferator-activated receptor beta.  
GN PPARBETA.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=20267232; PubMed=10809235;  
RA Takada I., Yu R.T., Xu H.E., Lambert M.H., Montana V.G., Kilewer S.A.,  
RA Evans R.M., Umeson K.,  
RT "Alteration of a single amino acid in peroxisome proliferator-  
RT activated receptor-alpha (PPAR alpha) generates a PPAR delta  
RT phenotype."  
RT Mol. Endocrinol. 14:733-740(2000).  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL; AF163810; AAF80480.1; -.  
DR HSSP; P03372; 1HCQ.  
DR InterPro: IPR000536; Hormone\_rec.1lg.  
DR InterPro: IPR001723; Sterm\_receptor.  
DR InterPro: IPR001628; Znf\_C4steroid.  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; Zf-C4; 1.  
DR PRINTS; PR00398; STRDHOMER.  
DR PRINTS; PR00047; STROIDFINGER.  
DR PRODOM; PD00035; Znf\_C4steroid; 1.  
DR SMART; SM00430; HOL1; 1.  
DR SMART; SM00399; ZNF\_C4; 1.  
DR PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
KW DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
SQ SEQUENCE 443 AA; 50108 MW; EB6E0C39554C76CD CQC64;

Query Match 31.9%; Score 1117.5; DB 13; Length 443;  
Best Local Similarity 53.2%; Pred. No. 2.4e-62;  
Matches 235; Conservative 68; Mismatches 88; Indels 51; Gaps 10;

OY 14 SDSFND--TISANISOEMTVDTMPFMPNFGISVDSLVMEDSHSFDIKPTTYDF 70  
DB 41 SSSYIDLSOSSSPSISDLOMGCEP-----TASGALNVECRVCGKASGF----- 85  
OY 71 SSISSPHY-----EDIP--FTRTDPVADYKYLKLOEQSAIKVEPASPYYSEKTOLY 123  
DB 86 -----HYGVHACGECGKGFRT-----IRMKL-EYEKER-----SKCIQK 121  
OY 124 NNNKQYGRFOGCLAVGSHNNAIRGCRMPQAEKELIAET-SSDIDQLNPESADIRALAK 182  
DB 122 NNNKQYGRFOGCLGSHNNAIRGCRMPQAEKELIAET-SSDIDQLNPESADIRALAK 181  
OY 183 HLYDSYIKSPFLTKAKRAIILTGK--TTDKSPVIYDMNSLMNGEDKIKFKHTPTLQDSK 241  
DB 182 HIYNAVILNFMNNTKKRAGKILTGKASSITPQPFVHDMITLMOAEGVLWKQIVNGITPYK 241  
OY 242 EVAIRIFOGCGRSVAVQETEVAKSIGFVNLNDQVTLTKYGVHEIITMLASLNN 301  
DB 242 EIGVHVFRCQCTVEYVRELTEFAKSIPIFGVLNDQVTLTKYGVHEIITMLASLNN 301  
OY 302 KDGVLISGCGMTREPLKSLKPPGDMERPEFAVKNALDSDSLAIFIAVILISG 361  
DB 302 KDGVLISGCGMTREPLKSLKPPGDMERPEFAVKNALDSDSLAIFIAVILISG 361

DB 302 KDGVLVANGNGVTRFEPLTKLPNEIMEPKFEFAVKNALDSDSLFVAAILIG 361  
OY 362 DRPGLLVNVPIDIDONDLQALELOLKLNHPESSOLFALOKMTDLROIVTEHVOV 421  
DB 362 DRPGLLVNVPIDIDONDLQALELOLKLNHPESSOLFALOKMTDLROIVTEHVOV 421  
OY 422 IKKTETMSLHPLQEIYKDL 443  
DB 422 IKKTETMSLHPLQEIYKDMY 443

## RESULT 14

OY95KZ4 PRELIMINARY: PRT: 219 AA.  
AC Q95KZ4;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Peroxisome proliferator-activated receptor gamma (Fragment).  
OS Mustela vison (American mink).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Carnivora; Fissipedia; Mustelidae; Mustelinae;  
OC Mustela.  
OX NCBI\_TaxID=9667;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX Desmarais J.A., Bennett R.D., Murphy B.D.;  
RA "Mustela vison Peroxisome Proliferator-Activated Receptor Gamma  
RT mRNA."  
RT Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTORS FAMILY.  
DR EMBL; AY050524; AAL1452.1; -.  
DR InterPro: IPR000536; Hormone\_rec.1lg.  
DR InterPro: IPR001628; Znf\_C4steroid.  
DR Pfam: PF00104; hormone\_rec; 1.  
DR Pfam: PF00105; Zf-C4; 1.  
DR DNA-binding; Nuclear protein; Receptor; Transcription regulation;  
KW Zinc-finger.  
FT NON\_TER 1 1  
FT NON\_TER 219 219  
SQ SEQUENCE 219 AA; 24828 MW; 8F1D1E4F90EA6320 CQC64;

Query Match 31.5%; Score 1105; DB 6; Length 219;  
Best Local Similarity 99.5%; Pred. No. 5.7e-62;  
Matches 218; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 136 CIAGVSHNNAIRFGMRPQAEKELIAETSSDIDQLNPESADIRALAKHLYDSYKSPFLT 195  
DB 1 CIAGVSHNNAIRFGMRPQAEKELIAETSSDIDQLNPESADIRALAKHLYDSYKSPFLT 60  
OY 196 KAKARAILTGKTTDKSPVIYDMNSLMNGEDKIKFKHTPTLQDSKQVAVIRFOGCGFRS 255  
DB 61 KAKARAILTGKTTDKSPVIYDMNSLMNGEDKIKFKHTPTLQDSKQVAVIRFOGCGFRS 120  
OY 256 VEAQVETEVAKSIGFVNLNDQVTLTKYGVHEIITMLASLNNKQVILISGCGMT 315  
DB 121 VEAQVETEVAKNIGFVNLNDQVTLTKYGVHEIITMLASLNNKQVILISGCGMT 180  
OY 316 REFLKSLRPPGDMERPEFAVKNALDSDSLAIFI 354  
DB 181 REFLKSLRPPGDMERPEFAVKNALDSDSLAIFI 219  
RESULT 15  
OY95N78 PRELIMINARY: PRT: 468 AA.  
AC Q95N78;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)  
DE Peroxisome proliferator activated receptor alpha.  
OS Canis familiaris (Dog).



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:23:21 ; Search time 20 seconds

(without alignments)  
1426.786 Million cell updates/sec

Title: US-09-931-007A-1

Perfect score: 3508

Sequence: 1 MGFTLGSDSPIDPESDSFTDT.....KTEIDMSLPLQLQEIYKDLV 688

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	505	1 PPAR_HUMAN	P37231 homo sapien
2	2216	63.2	505	1 PPAR_MACMO	O18824 macaca mula
3	2171.5	61.9	504	1 PPAR_PIG	O62807 sus scrofa
4	2153	61.4	505	1 PPAR_MOUSE	P37238 mus musculu
5	2148	61.2	505	1 PPAR_RAT	O88275 rattus norv
6	2132	60.8	505	1 PPAR_BOVIN	O18971 bos taurus
7	2043	58.2	475	1 PPAR_CRIGR	P57297 cricetus
8	2041	58.2	475	1 PPAR_RABIT	O19052 oryctolagus
9	1471	41.9	477	1 PPAR_XENLA	P37234 xenopus lae
10	1101.5	31.4	441	1 PPAR_XENLA	Q03181 homo sapien
11	1091.5	31.1	441	1 PPAR_RAT	P37230 rattus norv
12	1083.5	30.9	468	1 PPAR_MOUSE	P35396 mus musculu
13	1082.5	30.9	440	1 PPAR_MOUSE	P23204 mus musculu
14	1076.5	30.7	468	1 PPAR_HUMAN	Q07869 homo sapien
15	1060.5	30.2	474	1 PPAR_XENLA	P37232 xenopus lae
16	1042	29.7	467	1 PPAR_CAVPO	O35507 cavia porce
17	946.5	27.0	396	1 PPAR_XENLA	P37233 xenopus lae
18	431	12.3	606	1 E75_METEN	O7245 metapneus
19	408	11.6	576	1 NR2_MOUSE	O60674 mus musculu
20	404.5	11.5	579	1 NR2_MOUSE	O61495 homo sapien
21	397	11.3	578	1 NR2_MOUSE	O63504 rattus norv
22	392	11.2	1237	1 E75B_DROME	P17672 drosophila
23	385.5	11.0	1237	1 E75A_DROME	P17671 drosophila
24	379	10.8	614	1 NR1_HUMAN	P20393 homo sapien
25	375.5	10.7	459	1 NR1_HUMAN	P45446 rattus norv
26	371.5	10.6	508	1 NR1_RAT	O63503 rattus norv
27	371.5	10.6	1443	1 E75C_DROME	P13055 drosophila
28	365.5	10.4	459	1 NR1_HUMAN	O92753 homo sapien
29	354.5	10.1	711	1 E75_GALME	O92039 galliera me
30	351	10.0	699	1 E75_MANSE	O08893 manduca sex
31	344	9.8	690	1 E75_CHOFO	O01539 choriscoue
32	341.5	9.7	556	1 RORA_MOUSE	P35398 homo sapien
33	338.5	9.6	523	1 RORA_MOUSE	P51448 mus musculu

34	327.5	9.3	458	1 RRA_NOTVI	P18514 notophthalm
35	326.5	9.3	455	1 RRB_CHICK	P22448 gallus gall
36	326.5	9.3	455	1 RRB_CONJA	O946b3 coturnix co
37	326	9.3	447	1 RRA_PIGROU	O94523 fugu rubrip
38	325.5	9.3	560	1 RRB_HUMAN	P51449 homo sapien
39	324.5	9.3	455	1 RRB_HUMAN	P10826 homo sapien
40	323	9.2	865	1 E78A_DROME	P45447 drosophila
41	322	9.2	445	1 NR3_MOUSE	O920Y9 mus musculu
42	322	9.2	445	1 NR3_MOUSE	O62685 rattus norv
43	319.5	9.1	482	1 RRB_MOUSE	P22605 mus musculu
44	318	9.1	505	1 RRB_NOTVI	P18516 notophthalm
45	316	9.0	476	1 RRG1_XENLA	P51127 xenopus lae

## ALIGNMENTS

RESULT 1  
ID PPAR\_HUMAN STANDARD; PRT; 505 AA.  
AC P37231: Q14515; Q15178; Q15179; Q15832; Q00684; Q15180; Q00710;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 15-JUN-2002 (Rel. 41, Last annotation update)  
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).  
GN PPARC OR NR1C3.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_Taxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND CHARACTERIZATION.  
RC TISSUE=Heart;  
RX MEDLINE=97218249; PubMed=9065481;  
RA Mukherjee R., Jow L., Croston G.E., Paterniti J.R. Jr.;  
RT "Identification, characterization, and tissue distribution of human  
RT peroxisome proliferator-activated receptor (PPAR) isoforms PPARgamma2  
RT versus PPARgamma1 and activation with retinoid X receptor agonists and  
RT antagonists";  
RL J. Biol. Chem. 272:8071-8076(1997).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
RC TISSUE=Fat body;  
RX Elbrecht A., Chen Y., Cullinan C.A., Hayes N., Leibowitz M.D.,  
RA Moller D.E., Berger J.;  
RT "Molecular cloning, expression and characterization of human  
RT peroxisome proliferator activated receptor gamma 1 and gamma 2";  
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Adipose tissue;  
RX Kato S.;  
RL Submitted (FEB-1996) to the EMBL/Genbank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Bone marrow;  
RX MEDLINE=95307078; PubMed=7787419;  
RA Greene M.E., Blumberg B., McBride O.W., Yi H.F., Kronquist K.,  
RA Kwan K., Hsieh L., Greene G., Nimer S.D.;  
RT "Isolation of the human peroxisome proliferator activated receptor  
RT gamma cDNA: expression in hematopoietic cells and chromosomal  
RT mapping";  
RL Gene Expr. 4:281-299(1995).  
RN [5]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC TISSUE=Placenta;  
RX MEDLINE=98016122; PubMed=9356045;  
RA Okazawa H., Mori H., Tamori Y., Araki S., Niki T., Masugi J.,  
RA Kawachi M., Kubota T., Shinoda H., Kasuga M.;  
RT "No coding mutations are detected in the peroxisome proliferator-  
RT activated receptor-gamma gene in Japanese patients with lipotrophic  
RT diabetes";  
RL Diabetes 46:1904-1906(1997).

```

RN [6]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta:
RX MEDLINE=96305359; PubMed=8706692;
RA Lambe K.G., Tugwood J.D.;
RT "A human peroxisome-proliferator-activated receptor-gamma is
RT activated by inducers of adipogenesis, including thiazolidinedione
RT drugs.";
RL Eur. J. Biochem. 239:1-7(1996).
RN [7]
RP X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).
RX MEDLINE=20337987; PubMed=10882139;
RA Gampe R.T. Jr., Montana V.G., Lambert M.H., Miller A.B., Bledsoe R.K.,
RA Milburn M.V., Klierer S.A., Willson T.M., Xu H.E.;
RT "Asymmetry in the PARGgamma/RXRalpha crystal structure reveals the
RT molecular basis of heterodimerization among nuclear receptors.";
RL Mol. Cell 5:545-555(2000).
RN [8]
RP VARIANT ALA-12.
RX MEDLINE=98086341; PubMed=9425261;
RA Yen C.-J., Beamer B.A., Negri C., Silver K., Brown K.A., Yarnall D.P.,
RA Burns D.K., Roth J., Shuldiner A.R.;
RT "Molecular scanning of the human peroxisome proliferator activated
RT receptor gamma (hPPAR-gamma) gene in diabetic Caucasians:
RT identification of a pro12ala PPAR-gamma-2 missense mutation.";
RL Biochem. Biophys. Res. Commun. 241:270-274(1997).
RN [9]
RP VARIANT OBESITY GLN-113.
RX MEDLINE=98418646; PubMed=97537710;
RA Ristow M., Muller-Wieland D., Pfeiffer A., Krone W., Kahn C.R.;
RT "Obesity associated with a mutation in a genetic regulator of
RT adipocyte differentiation.";
RL New Engl. J. Med. 339:953-959(1998).
RN [10]
RP VARIANT ALA-12.
RX MEDLINE=99337654; PubMed=10407229;
RA Hamann A., Munzberg H., Button P., Busing B., Hinney A., Mayer H.,
RA Siegfried W., Hebebrand J., Grotten H.;
RT "Missense variants in the human peroxisome proliferator-activated
RT receptor-gamma2 gene in lean and obese subjects.";
RL Eur. J. Endocrinol. 141:90-92(1999).
RN [11]
RP VARIANTS COLON CANCER PRO-314 AND HIS-316, AND VARIANT ALA-12.
RX MEDLINE=99322672; PubMed=10394368;
RA Sarraf P., Mueller E., Smith W.W., Wright H.M., Kum J.B.,
RA Aaltonen L.A., de la Chapelle A., Spiegelman B.M., Eng C.;
RT "Loss-of-function mutations in PPAR-gamma associated with human colon
RT cancer.";
RL Mol. Biol. Cell 3:799-804(1999).
RN [12]
RP VARIANTS DIABETES MET-318 AND LEU-495.
RX MEDLINE=20085964; PubMed=10622252;
RA Barroso I., Gurnell M., Crowley V.E.F., Agostini M., Schwabel J.W.,
RA Soos M.A., Masien G.L., Williams T.D.M., Lewis H., Schaefer A.J.,
RA Chatterjee V.K.K., O'Rahilly S.;
RT "Dominant negative mutations in human PPAR-gamma associated with
RT severe insulin resistance, diabetes mellitus and hypertension.";
RL Nature 402:880-883(1999).
RN -1-
RP FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
RP HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
RP THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
RP OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
RP PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
RP OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN
CC SKELETAL MUSCLE, SPLEEN, HEART AND LIVER. ALSO ARE DETECTABLE IN
CC PLACENTA, LUNG AND OVARY.
CC -1- DISEASE: DEFECTS IN PARG CAN LEAD TO TYPE 2 INSULIN-RESISTANT
CC DIABETES AND HYPERTENSION.
CC -1- DISEASE: DEFECTS IN PARG COULD PLAY A ROLE IN THE GENETIC
CC PREDISPOSITION TO OBESITY.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U79012; AAC51248.1; -
DR EMBL: U63415; AAB04028.1; -
DR EMBL: D83233; BAA18949.1; -
DR EMBL: LA0904; AAA80314.1; -
DR EMBL: AB005526; BAA23354.1; ALT_INIT.
DR EMBL: AB005521; BAA23354.1; JOINED.
DR EMBL: AB005522; BAA23354.1; JOINED.
DR EMBL: AB005523; BAA23354.1; JOINED.
DR EMBL: AB005524; BAA23354.1; JOINED.
DR EMBL: AB005525; BAA23354.1; JOINED.
DR EMBL: X90563; CAA62152.1; ALT_INIT.
DR EMBL: X90563; CAA62153.1; -
DR PIR: S42489; S42489.
DR PDB: 1FM6; 16-FEB-01.
DR PDB: 1FM9; 16-FEB-01.
DR TRANSFAC: T03731; -
DR Genew: HGNC:9236; PARG.
DR MIM: 601487; -
DR MIM: 604367; -
DR InterPro: IPR000536; Hormone_rec_1lg.
DR InterPro: IPR001723; Stdhrrm_receptor.
DR InterPro: IPR001628; ZnI_C4steroid.
DR Pfam: PF00105; hormone_rec_1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRODHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; ZnI_C4steroid; 1.
DR SMART: SM00430; HOI1; 1.
DR SMART: SM00399; ZnI_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor: Transcription regulation; Activator: DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Alternative splicing;
KW Phosphorylation; Polymorphism; Disease mutation; Diabetes mellitus;
KW Obesity; 3D-structure.
FT DNA_BIND 139 203
FT ZN_FING 139 159
FT ZN_FING 176 198
FT DOMAIN 318 505
FT MOD_RES 112 112
FT VARSPPLIC 1 30
FT VARIANT 12 12
FT VARIANT 113 113
FT VARIANT 314 314
FT VARIANT 316 316
FT VARIANT 318 318
FT VARIANT 495 495
FT CONFLICT 36 37
FT CONFLICT 213 214
FT CONFLICT 240 240
FT CONFLICT 424 426
FT SEQUENCE 505 AA; 57620 MW; 3933EF36A0ECAFC64;

```

Query Match 63.6%; Score 2231; DB 1; Length 505;  
 Best Local Similarity 87.7%; Pred. No. 1.7e-116;  
 Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 1 MGTTLDSPIDPSDFSTLTLSANISQEMTVDTEPMPPTNFGISSVDLSVMDHSHSF 60  
 DB 1 MGTTLDSPIDPSDFSTLTLSANISQEMTVDTEPMPPTNFGISSVDLSVMDHSHSF 60  
 QY 61 DLRPTTVDFSSISPHYEDIPFTPTDPAVDYKYDKLQEOSAKVBPASPPYSEKT 120  
 DB 61 DLRPTTVDFSSISPHYEDIPFTPTDPAVDYKYDKLQEOSAKVBPASPPYSEKT 120  
 QY 121 QLYN----- 124  
 DB 121 QLYNKPHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRTRILKLIYDRCDLNC 180  
 QY 125 -----RNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLQNPESADLR 178  
 DB 181 RHKRSRNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLQNPESADLR 240  
 QY 179 ALAKHLVDSYISFPLTKAKARALLTGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 238  
 DB 241 ALAKHLVDSYISFPLTKAKARALLTGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 300  
 QY 239 OSKEVAIRIFOGQCRFSVAVOEITEYAKSIPGFVNLINDQVTLTKYGVHEIITMLAS 298  
 DB 301 OSKEVAIRIFOGQCRFSVAVOEITEYAKSIPGFVNLINDQVTLTKYGVHEIITMLAS 360  
 QY 299 LMKKGVLSISGQGFMTREFLSLRKPPGDEMEKPEFPAVKFNALELDDSDLAIFAVIY 358  
 DB 361 LMKKGVLSISGQGFMTREFLSLRKPPGDEMEKPEFPAVKFNALELDDSDLAIFAVIY 420  
 QY 359 LSGDRGGLNVPRIEDIONLQALELOLKLNPRESSOLFAPKLQKMTDLROIVHEVOL 418  
 DB 421 LSGDRGGLNVPRIEDIONLQALELOLKLNPRESSOLFAPKLQKMTDLROIVHEVOL 480  
 QY 419 LOVIRKRTDMSLHPLOEIRKDY 443  
 DB 481 LOVIRKRTDMSLHPLOEIRKDY 505

RESULT 2  
 PRAT\_MACMU STANDARD; PRT; 505 AA.  
 AC 018924; Q9TQW6;  
 DT 16-OCT-2001 (rel. 40, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).  
 GN PPARC OR NR1C3.  
 OS Macaca mulatta (Rhesus macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9544;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).  
 RC TISSUE=adipose tissue;  
 RX MEDLINE=99021153; PubMed=9806316;  
 RA Hotta K., Gustafson T.A., Yoshikawa S., Ortmeyer H.K., Bodkin N.L.,  
 RA Hansen B.C.;  
 RT "Relationships of PPARgamma and PPARgamma2 mRNA levels to obesity,  
 RT diabetes and hyperinsulinemia in rhesus monkeys";  
 RL Int. J. Obes. Relat. Metab. Disord. 22:1000-1010(1998).  
 CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS  
 CC HYPOGLYCEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR  
 CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.  
 CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.

CC -!- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);  
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.  
 CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN  
 CC LIVER, HEART, KIDNEY, STOMACH, DUODENUM AND COLON.  
 CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 CC NR1 SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF033103; AAB87480.1; -;  
 CC EMBL; AF033343; AAB87482.1; -;  
 CC EMBL; AF033342; AAB87481.1; -;  
 CC HSSP; P37231; 1FM9.  
 CC InterPro: IPR000536; Hormone\_rec.1lg.  
 CC InterPro: IPR001723; sthrrm\_receptor.  
 CC InterPro: IPR001628; znf\_C4steroid.  
 CC Pfam; PF00104; hormone\_rec.1.  
 CC Pfam; PF00105; zf-C4; 1.  
 CC PRINTS; PR00398; STRDHOMER.  
 CC PRINTS; PR00047; STROIDINGER.  
 CC Prodom; PD000035; znf\_C4steroid; 1.  
 CC SMART; SM00430; HOLY; 1.  
 CC SMART; SM00399; znf\_C4; 1.  
 CC PROSITE; PS00031; NUCLEAR\_RECEPTOR; 1.  
 CC Receptor; Transcription regulation; Activator; DNA-binding;  
 CC Nuclear protein; Zinc-finger; Multigene family; Alternative splicing;  
 CC Phosphorylation.  
 CC DNA\_BIND 139 203 NUCLEAR RECEPTOR-TYPE.  
 CC ZN\_FING 139 159 C4-TYPE.  
 CC ZN\_FING 176 198 C4-TYPE.  
 CC DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).  
 CC MOD\_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY  
 CC SIMILARITY).  
 CC VARSPLIC 1 30 MISSING (IN ISOFORM 1).  
 CC SEQUENCE 505 AA; 57590 MW; 41836a624AAAF942 CRC64;

Query Match 63.2%; Score 2216; DB 1; Length 505;  
 Best Local Similarity 86.9%; Pred. No. 1.1e-115;  
 Matches 439; Conservative 3; Mismatches 1; Indels 62; Gaps 1;

QY 1 MGTTLDSPIDPSDFSTLTLSANISQEMTVDTEPMPPTNFGISSVDLSVMDHSHSF 60  
 DB 1 MGTTLDSPIDPSDFSTLTLSANISQEMTVDTEPMPPTNFGISSVDLSVMDHSHSF 60  
 QY 61 DLRPTTVDFSSISPHYEDIPFTPTDPAVDYKYDKLQEOSAKVBPASPPYSEKT 120  
 DB 61 DLRPTTVDFSSISPHYEDIPFTPTDPAVDYKYDKLQEOSAKVBPASPPYSEKT 120  
 QY 121 QLYN----- 124  
 DB 121 QLYNKPHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRTRILKLIYDRCDLNC 180  
 QY 125 -----RNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLQNPESADLR 178  
 DB 181 RHKRSRNKCQYCRFOCKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDLQNPESADLR 240  
 QY 179 ALAKHLVDSYISFPLTKAKARALLTGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 238  
 DB 241 ALAKHLVDSYISFPLTKAKARALLTGKTTDKSPFYIYDMNSLMGEGDKIRKHTHPLOE 300  
 QY 239 OSKEVAIRIFOGQCRFSVAVOEITEYAKSIPGFVNLINDQVTLTKYGVHEIITMLAS 298  
 DB 301 OSKEVAIRIFOGQCRFSVAVOEITEYAKSIPGFVNLINDQVTLTKYGVHEIITMLAS 360  
 QY 299 LMKKGVLSISGQGFMTREFLSLRKPPGDEMEKPEFPAVKFNALELDDSDLAIFAVIY 358  
 DB 361 LMKKGVLSISGQGFMTREFLSLRKPPGDEMEKPEFPAVKFNALELDDSDLAIFAVIY 420

```

OY 359 LSGDRPGLNVKPIEDIDONLLOALELQKLNHPRESSOLFAPKLLQKMTDLROIVTEHVOL 418
DB 421 LSGDRGLNVKPIEDIDONLLOALELQKLNHPRESSOLFAPKLLQKMTDLROIVTEHVOL 480
OY 419 LOYIKKETDMSLHPLLOEITYKDL 443
DB 481 LOYIKKETDMSLHPLLOEITYKDL 505

RESULT 3
PRT_PIG STANDARD: PRT; 504 AA.
ID PPAT_PIG
AC 062807; 077815;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARC OR NR1C3.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC TISSUE=adipose tissue;
RC MEDLINE=99132654; PubMed=991452;
RA Houseknecht K.L., Bidwell C.A., Portocarrero C.P., Spurlock M.E.;
RT "Expression and cDNA cloning of porcine peroxisome proliferator-
RT activated receptor gamma (PPARgamma).";
RL Gene 225:89-96(1998).
RN [2]
RN SEQUENCE FROM N.A. (ISOFORM 2).
RX MEDLINE=99410873; PubMed=10481259;
RA Ding S.T., McNeil R.L., Mersmann H.J.;
RT "Expression of porcine adipocyte transcripts: tissue distribution and
RT differentiation in vitro and in vivo.";
RL Comp. Biochem. Physiol. 123B:307-318(1999).
RN [3]
RN SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RP STRAIN=Duroc, and Norwegian Landrace; TISSUE=adipocyte;
RX MEDLINE=98401156; PubMed=9731203;
RA Grindler E., Sundvold H., Klungland H., Lien S.;
RT "Characterisation of porcine peroxisome proliferator-activated
RT receptors gamma 1 and gamma 2: detection of breed and age differences
RT in gene expression.";
RL Biochem. Biophys. Res. Commun. 249:713-718(1998).
RL
CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 1 AND 2 (SHOWN HERE); ARE
CC PRODUCED BY ALTERNATIVE SPLICING.
CC -!- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE AND LOWER
CC IN SPLEEN. VERY LOW LEVELS IN KIDNEY, INTESTINE, LUNG AND MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF059245; AAC14348.1; -
CC EMBL: AF103946; AAD19577.1; -
CC EMBL: AJ006757; CAA07225.1; -

```

```

DR EMBL: AJ006756; CAA07224.1; -.
DR HSSP: P37231; 1FM9.
DR InterPro: IPR000536; Hormone_rec_1lg.
DR InterPro: IPR001723; Stdhm_receptor.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4_1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; ZnF_C4steroid; 1.
DR SMART: SM00430; HOL1.1.
DR SMART: SM00399; ZnF_C4_1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
DR Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; zinc-finger; Multigene family; Alternative splicing;
KW Phosphorylation.
FT DNA_BIND 138 202 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 138 158 C4-TYPE.
FT ZN_FING 175 197 C4-TYPE.
FT DOMAIN 317 504 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 111 111 PHOSPHORYLATION (BY MAPK) (BY
FT VARSPLIT 1 29 MISSING (IN ISOFORM 1).
FT CONFLICT 381 381 K -> R (IN REF. 3).
FT CONFLICT 426 426 G -> R (IN REF. 3).
SQ SEQUENCE 504 AA; 57512 MW; 57C0BDE901493D31 CRC64;

Query Match 61.9%; Score 2171.5; DB 1; Length 504;
Best Local Similarity 85.5%; Pred. No. 3,2e-113;
Matches 432; Conservative 6; Mismatches 4; Indels 63; Gaps 2;

OY 1 MGFTLSDSPIDPSDFTTSLANISQEMTWDTMPFWPTNGISSVLSVMEHSHSF 60
DB 1 MGFTLSDSLIDPSDAF-DTSLANISQEVMTWDTMPFWPTNGISSVLSVWDHSHSF 59
OY 61 DIKFTTVPDSISTPHYEDIPTRDPVADYKYDLQETOSAIVKVPASPPYSEKT 120
DB 60 DIKFTTVPDSISTPHYEDIPTRDPVADYKYDLQDYOSAIVKVPASPPYSEKT 119
OY 121 QLVN----- 124
DB 120 QLVNKHHEPNSLMAIECRVCGDKASGFHYGVHACGCKGFFRRTRRLKIYDRCDLNC 179
OY 125 -----RNKCOYCRFOKCAVGNSHNAIRFGMPQAEKELAEISSDIDOLNPESADLR 178
DB 180 RHKKSRNKCQYCRFOKCAVGNSHNAIRFGMPQAEKELAEISSDIDOLNPESADLR 239
OY 179 ALAKHLVDSYISFPLTKAKARAILTGKTTDKSPVIYVYMNLSLMDGEDIKRNHTPLOE 238
DB 240 ALAKHLVDSYISFPLTKAKARAILTGKTTDKSPVIYVYMNLSLMDGEDIKRNHTPLOE 299
OY 239 QSKVEAIRIFOGQFVSVAVOEITFYAKSIGFVLDLNDQVTLKYVNHETIYMLAS 298
DB 300 QSKVEAIRIFOGQFVSVAVOEITFYAKNIGFVLDLNDQVTLKYVNHETIYMLAS 359
OY 299 LMKKDVLLSEGGFMTREFLSLRKPFQDFMEPKREFAVKFNALDELSDSLAIFAVYII 358
DB 360 LMKKDVLLSEGGFMTREFLSLRKPFQDFMEPKREFAVKFNALDELSDSLAIFAVYII 419
OY 419 LSGDRPGLNVKPIEDIDONLLOALELQKLNHPRESSOLFAPKLLQKMTDLROIVTEHVOL 418
DB 420 LSGDRPGLNVKPIEDIDONLLOALELQKLNHPRESSOLFAPKLLQKMTDLROIVTEHVOL 479
OY 419 LOYIKKETDMSLHPLLOEITYKDL 443
DB 480 LOYIKKETDMSLHPLLOEITYKDL 504

RESULT 4
PRT_MOUSE STANDARD: PRT; 505 AA.
ID PPAT_MOUSE
AC P37238;
DT 01-OCT-1994 (Rel. 30, Created)

```



DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).  
 OS PPARG OR NR1C3.  
 GN Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID:10090;  
 RN [1]  
 RN [1] SEQUENCE FROM N.A. (ISOFORM 2).  
 RC TISSUE=adipose tissue;  
 RX MEDLINE=95011536; PubMed=7926726;  
 RA Tontonoz P., Hu E., Graves R.A., Budavari A.I., Spiegelman B.M.;  
 RT "mPPAR gamma 2: tissue-specific regulator of an adipocyte enhancer.";  
 RL Gene Dev. 8:1224-1234(1994).  
 RN [2]  
 RN [2] SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=BALE/c; TISSUE=Heart;  
 RX MEDLINE=94059089; PubMed=8240342;  
 RA Chen F., Law S.W., O'Malley B.W.;  
 RT "Identification of two mPPAR related receptors and evidence for the  
 RT existence of five subfamily members.";  
 RL Biochem. Biophys. Res. Commun. 196:671-677(1993).  
 RN [3]  
 RN [3] SEQUENCE FROM N.A. (ISOFORM 1).  
 RC STRAIN=C57BL/6 X CBA; TISSUE=Liver;  
 RX MEDLINE=94086482; PubMed=8262913;  
 RA Zhu Y., Alvares K., Huang Q., Rao M.S., Reddy J.K.;  
 RT "Cloning of a new member of the peroxisome proliferator-activated  
 RT receptor gene family from mouse liver.";  
 RL J. Biol. Chem. 268:26817-26820(1993).  
 RN [4]  
 RN [4] SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Liver;  
 RX MEDLINE=94316694; PubMed=8041794;  
 RA Klierer S.A., Forman B.M., Blumberg B., Ong E.S., Borgmeyer U.,  
 RA Mangelsdorf D.J., Umesono K., Evans R.M.;  
 RT "Differential expression and activation of a family of murine  
 RT peroxisome proliferator-activated receptors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:7355-7359(1994).  
 RN [5]  
 RN [5] SEQUENCE FROM N.A.  
 RX MEDLINE=96249427; PubMed=8647948;  
 RA Vidal-Puig A., Jimenez-Linan M., Lowell B.B., Hamann A., Hu E.,  
 RA Spiegelman B., Flier J.S., Moller D.E.;  
 RT "Regulation of PPAR gamma gene expression by nutrition and obesity in  
 RT rodents.";  
 RL J. Clin. Invest. 97:2553-2561(1996).  
 CC -i- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS  
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR  
 CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.  
 CC -i- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
 CC -i- SUBCELLULAR LOCATION: Nuclear.  
 CC -i- ALTERNATIVE PRODUCTS: 2 ISOFORMS; 1 AND 2 (SHOWN HERE); ARE  
 CC PRODUCED BY ALTERNATIVE SPLICING.  
 CC -i- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. ALSO  
 CC FOUND IN LIVER, SKELETAL MUSCLE, HEART, ADRENAL GLAND, SPLEEN,  
 CC KIDNEY AND INTESTINE.  
 CC -i- DEVELOPMENTAL STAGE: IT APPEARS FIRST AT DAY 13.5 POSTCONCEPTION,  
 CC AND INCREASES UNTIL BIRTH.  
 CC -i- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 CC NRI SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL: U09138; AAA62277.1; -  
 DR EMBL: U01664; AAA62110.1; -  
 DR EMBL: U01841; AAC52134.1; -  
 DR EMBL: U10374; AAA19971.1; -  
 DR PIR: JN0881; JN0881.  
 DR HSSP: P37231; 1FM9.  
 DR TRANSFAC: T02529; -  
 DR MGD: MGI:97747; Pradp.  
 DR InterPro: IPR000536; Hormone\_rec\_1lg.  
 DR InterPro: IPR001723; Stohrm\_recceptor.  
 DR InterPro: IPR001628; Znf\_C4steroid.  
 DR Pfam: PF00104; hormone\_rec.1.  
 DR Pfam: PF00105; zf-C4.1.  
 DR PRINTS: PR00398; STRDHORMONER.  
 DR PRINTS: PR00047; STROIDFINGER.  
 DR PRODOM: PD000035; Znf\_C4steroid.1.  
 DR SMART: SM00430; HOL1.1.  
 DR SMART: SM00399; Znf\_C4.1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR.1.  
 KW Receptor; Transcription regulation; Activator; DNA-binding;  
 KW Nuclear protein; zinc-finger; Multigene family; Alternative splicing;  
 KW Phosphorylation.  
 KW DNA\_BIND 139 203 NUCLEAR RECEPTOR-TYPE.  
 FT ZN\_FING 139 159 C4-TYPE.  
 FT ZN\_FING 176 198 C4-TYPE.  
 FT DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).  
 FT MOD\_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY  
 FT SIMILARITY).  
 FT VARSPLIC 1 30 MISSING (IN ISOFORM 1).  
 FT CONFLICT 213 214 MP -> DR (IN REF. 2).  
 FT CONFLICT 281 283 NSL -> SSF (IN REF. 2).  
 FT CONFLICT 383 383 N -> S (IN REF. 2 AND 4).  
 FT CONFLICT 497 497 L -> F (IN REF. 2).  
 SQ SEQUENCE 505 AA; 57598 MW; AB8F3F6086E2A10A CRC64;  
 Query Match 61.4%; Score 2153; DB 1; Length 505;  
 Best Local Similarity 84.0%; Pred. No. 3.4e-112;  
 Matches 424; Conservative 9; Mismatches 10; Indels 62; Gaps 1;  
 QY 1 MGETLDSPTIDESDFTTLTSLANSIQEMTVNDEMPNPNFGISSVDSLWEDHSSEF 60  
 DB 1 METLDSVDDEHGAFAALPMSTQETITVNDTEMPNPNFGISSVDSLWEDHSSEF 60  
 QY 61 DLRPFTVFSSTSPHYEDIPFTPTPVADYKYDKLQEOYSAIKVPASPPYSEKT 120  
 DB 61 DLRPFTVFSSTSPHYEDIPFTPTPVADYKYDKLQEOYSAIKVPASPPYSEKT 120  
 QY 121 QLNN----- 124  
 DB 121 QLNNRHEPNSLMAIECRVCGDKASGFHYVHACGCKGFFRTIRLKLIRCDLNC 180  
 QY 125 -----RNCQYCRFOKCLAVGMSHNAIRFGMPQAEKRLAEISSDIDOLNPESADLR 178  
 DB 181 RIHKRSRNCQYCRFOKCLAVGMSHNAIRFGMPQAEKRLAEISSDIDOLNPESADLR 240  
 QY 179 ALAKHLVDSYISFPLTKAKARAILGKTTDSPFTYVMNSLMGEGDKIRKHTIPLOE 238  
 DB 241 ALAKHLVDSYISFPLTKAKARAILGKTTDSPFTYVMNSLMGEGDKIRKHTIPLOE 300  
 QY 239 QSEVAIRIFQCGPFSVAVOEITEYAKSIGFVNLNDQVTLTKYGVHEITVTLAS 298  
 DB 301 QSEVAIRIFQCGPFSVAVOEITEYAKSIGFVNLNDQVTLTKYGVHEITVTLAS 360  
 QY 299 LNNKGVLLISEQGFTRFELSLKRPFGDEMPKFEFVKNALBLDSDLAIFLAVII 358  
 DB 361 LNNKGVLLISEQGFTRFELSLKRPFGDEMPKFEFVKNALBLDSDLAIFLAVII 420  
 QY 359 LSGDRGLNVKPIEDIONLQALELQKLNHPSSOLFALLOKMPDLROIVTEHVOL 418  
 DB 421 LSGDRGLNVKPIEDIONLQALELQKLNHPSSOLFALLOKMPDLROIVTEHVOL 480  
 QY 419 LQVIRKTTETDMSLHPLQEIYKDL 443

```

Db      481 LHVIRKTERDMSLHPLQEIYKNDLY 505
      1 |||
RESULT 5
AC      088275; O9QWGO: G9R197;
ID      PPAR_RAT STANDARD; PRT; 505 AA.
DT      16-OCT-2001 (Rel. 40, Created)
DT      16-OCT-2001 (Rel. 40, Last sequence update)
DT      15-JUN-2002 (Rel. 41, Last annotation update)
DE      Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN      PPARG OR NR1C3.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      [1]
RP      SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC      STRAIN=Sprague-Dawley; TISSUE=Brown adipose tissue;
RX      MEDLINE=99367468; PubMed=10438514;
RA      Guardiola-Diaz H.M., Rehmark S., Usuda N., Albrektsson T.,
RA      Felkamp D., Gustafsson J.-A., Alexson S.E.H.;
RT      "Rat peroxisome proliferator-activated receptors and brown adipose
RL      tissue function during cold acclimatization.";
RN      J. Biol. Chem. 274:23368-23377(1999).
RN      [2]
RP      SEQUENCE FROM N.A. (ISOFORM 2).
RC      STRAIN=Sprague-Dawley; TISSUE=adipose tissue;
RA      Tanaka T., Itoh H.;
RT      "Down-regulation of PPAR gamma.";
RL      Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN      [3]
RP      SEQUENCE FROM N.A. (ISOFORM 2).
RC      Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN      [4]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RC      TISSUE=adipocyte;
RA      Miyakita A., Okuno S., Watanabe T.K., Oga K., Tsuji A., Hishigaki H.,
RA      Suto T., Nakagawa K., Nakahara Y., Higashi K.;
RT      "Molecular cloning of rat PPAR-gamma gene.";
RL      Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A. (ISOFORM 1).
RC      STRAIN=Long Evans;
RX      MEDLINE=20259576; PubMed=10797535;
RA      Ershov A.V., Bazan N.G.;
RT      "Photoreceptor phagocytosis selectively activates PPARgamma expression
RL      in retinal pigment epithelial cells.";
RN      J. Neurosci. Res. 60:328-337(2000).
RN      [6]
RP      PHOSPHORYLATION.
RX      MEDLINE=97113205; PubMed=8953045;
RA      Hu E., Kim J.B., Sarraf P., Spiegelman B.M.;
RT      "Inhibition of adipogenesis through MAP kinase-mediated
RL      phosphorylation of PPARgamma.";
RN      Science 274:2100-2103(1996).
RN      [7]
RP      PHOSPHORYLATION OF SER-112, AND MUTAGENESIS OF SER-112.
RX      MEDLINE=97184167; PubMed=9030579;
RA      Adams M., Regnato M.J., Shao D., Lazar M.A., Chatterjee V.K.;
RT      "Transcriptional activation by peroxisome proliferator-activated
RL      receptor gamma is inhibited by phosphorylation at a consensus
RT      mitogen-activated protein kinase site.";
RN      J. Biol. Chem. 272:5128-5132(1997).
RN      [8]
RP      PHOSPHORYLATION.
RX      MEDLINE=97112959; PubMed=8943212;
RA      Zhang B., Berger J., Zhou G., Elbrecht A., Biswas S.,
RA      White-Carrington S., Szalkowski D., Moller D.E.;
RT      "Insulin- and mitogen-activated protein kinase-mediated
RT      phosphorylation and activation of peroxisome proliferator-activated

```

```

RT      receptor gamma.";
RL      J. Biol. Chem. 271:31771-31774(1996).
CC      - FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC      HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC      THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC      OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC      PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC      OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC      - SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC      - SUBCELLULAR LOCATION: Nuclear.
CC      - ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
CC      ARE PRODUCED BY ALTERNATIVE SPLICING.
CC      - TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE.
CC      - PTM: PHOSPHORYLATED BY MAPK. THE PHOSPHORYLATION INHIBITS
CC      PPAR GAMMA ACTIVITY.
CC      - SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC      NR1 SUBFAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC      or send an email to license@sib-sib.ch).
CC      -----
DR      EMBL; AF156666; AAD40119.1; -
DR      EMBL; AF156665; AAD40118.1; -
DR      EMBL; AB019561; BAA36485.1; -
DR      EMBL; Y12882; CAA73382.2; -
DR      EMBL; AB011365; BAA32540.1; -
DR      EMBL; AF246457; AAF63385.1; -
DR      EMBL; AF246458; AAF63386.1; -
DR      HSSP; P37231; 1FM9.
DR      InterPro: IPR000536; Hormone_rec_1lg.
DR      InterPro: IPR001723; Stdhrm_receptor.
DR      InterPro: IPR001628; ZnF_C4steroid.
DR      Pfam; PF00104; hormone_rec; 1.
DR      Pfam; PF00105; zf-C4; 1.
DR      PRINTS; PR00398; STROLDHOMNER.
DR      PRINTS; PR00047; STROLDHOMNER.
DR      PRODOM; PD000035; ZnF_C4steroid; 1.
DR      SMART; SM00430; HOL1; 1.
DR      SMART; SM00399; ZnF_C4; 1.
DR      PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW      Receptor; Transcription regulation; Activator; DNA-binding;
KW      Nuclear protein; zinc-finger; Multigene family; Alternative splicing;
KW      Phosphorylation.
FT      DNA_BIND 139..203 NUCLEAR RECEPTOR-TYPE.
FT      ZN_FING 139..159 C4-TYPE.
FT      ZN_FING 176..198 C4-TYPE.
FT      DOMAIN 318..505 LIGAND-BINDING (POTENTIAL).
FT      MOD_RES 112..112 PHOSPHORYLATION (BY MAPK) (BY
FT      VARSPLIC 1 30 SIMILARITY).
FT      MUTAGEN 112 112 MISSING (IN ISOFORM 1).
FT      CONFLICT 111 111 S->A: INCREASES ADIPOGENIC ACTIVITY.
FT      CONFLICT 111 111 A -> R (IN REF. 3).
SQ      SEQUENCE 505 AA; 57567 MW; F16E5CAB122EBB32 CRC64;
Query Match 61.2%; Score 2148; DB 1; Length 505;
Best Local Similarity 83.8%; Pred. No. 6,4e-112;
Matches 423; Conservative 9; Mismatches 11; Indels 62; Gaps 1;
QY      1 MGGTLDSPIDPESDSEFTPTLSANISQEMTWDTMEPWTNGISSVLDVSDHSHSF 60
DB      1 MGGTLDPPVDPEHGAFALPMPSTSQEITWDTMEPWTNGISSVLDVSDHSHSF 60
QY      61 DIKFTTVDFSSISPHYEDIPFTRTDPVADYKYDKLQEOSAIKVPASPPYSEKT 120
DB      61 DIKFTTVDFSSISAHYEDIPFTRTADPVADYKYDKLQEOSAIKVPASPPYSEKT 120
QY      121 QLVN----- 124

```

```

Db 121 OLYNRPHEPSNSLMAIECRVCGDKASGFHYHACBCKGFFRTIRLKLITYRCDLNC 180
QY 125 -----NRKCCOYCRFOCKLAVGSMHNATRFGRMPQAEKEKLLAETSSIDDLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOCKLAVGSMHNATRFGRMPQAEKEKLLAETSSIDDLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHITPLOE 238
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHITPLOE 300
QY 239 OSKEVAIRIFOGCCFRSVEAVOEITEYAKSIPGFVNLDLNOVTLTKGVHEIITMLAS 298
Db 301 OSKEVAIRIFOGCCFRSVEAVOEITEYAKNIPGFVNLNDLNOVTLTKGVHEIITMLAS 360
QY 299 LMNKDGVLSIEGCGFMTEFLSKLRPGDMEPKFEFAVKNALDELDDSLAIFIAVII 358
Db 361 LMNKDGVLSIEGCGFMTEFLSKLRPGDMEPKFEFAVKNALDELDDSLAIFIAVII 420
QY 359 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLQKMTDLRQIYTEHVOL 418
Db 421 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLQKMTDLRQIYTEHVOL 480
QY 419 LQVIKKTETDMSLHPLLOEIKDLY 443
Db 481 LQVIKKTETDMSLHPLLOEIKDLY 505

```

## RESULT 6

```

PPAT_BOVIN
ID PPAT_BOVIN STANDARD; PRT; 505 AA.
AC 018971;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPAR OR NR1C3.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;

```

```

RN [1]
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RT TISSUE=Fat;

```

```

RX MEDLINE=98042483; PubMed=9367859;
RA Sundvold H., Brzozowska A., Lien S.;

```

```

RT "Characterisation of bovine peroxisome proliferator-activated
RT receptors gamma 1 and gamma 2: genetic mapping and differential
RT expression of the two isoforms.";
RL Biochem. Biophys. Res. Commun. 239:857-861(1997).

```

```

CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PROXIMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS; 1 AND 2 (SHOWN HERE);
CC ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN ADIPOSE TISSUE. LOWER IN
CC SPLEEN AND LUNG. ALSO DETECTED IN OVARY MAMMARY GLAND AND SMALL
CC INTESTINE.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.

```

```

CC NRI SUBFAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
DR EMBL: Y12420; CAAT3033.1;
DR EMBL: Y12419; CAAT3032.1;
DR HSSP: P37231; 1FM9.
DR TRANSFAC: T04780;
DR InterPro: IPR000536; Hormone_rec_1lg.
DR InterPro: IPR001723; Slc4hrm_receptor.
DR InterPro: IPR001628; ZnF_C4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHORMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; ZnF_C4steroid; 1.
DR SMART: SM00430; HOLT; 1.
DR SMART: SM00399; znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Alternative splicing;
KW Phosphorylation.
FT DNA_BIND 139 203 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 139 159 C4-TYPE.
FT ZN_FING 176 198 C4-TYPE.
FT DOMAIN 318 505 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 112 112 PHOSPHORYLATION (BY MAPK) (BY
FT SIMILARITY).
FT VARSPLIC 1 30 MISSING (IN ISOFORM 1).
FT SEQUENCE 505 AA; 57579 MW; 5F20B115087B3C83 CRC64;

```

```

Query Match 60.8%; Score 2132; DB 1; Length 505;
Best Local Similarity 83.6%; Pred. No. 4,9e-111;
Matches 422; Conservative 11; Mismatches 10; Indels 62; Gaps 1;

```

```

QY 1 MGEITGDSPTIDPESDSEFDPLSANTISOEMTVDMPMPNPFGISVDLSVMEDHSF 60
Db 1 MGEITGDLIDPESDSEFPVIVSARTSQEITVDTMPMPNPFGISVDLSMMDHSHAF 60
QY 61 DIKPFPTVDFSSISIPRHEDIPFTRTDPVADVADKYDKLQEQSAIKKEPAPPYSEKT 120
Db 61 DIKPFPTVDFSSISIPRHEDIPFPADPMVADKYDKLQEQSAIKKEPSPYISKKT 120
QY 121 QLY
Db 121 QLY
QY 124 -----NRKCCOYCRFOCKLAVGSMHNATRFGRMPQAEKEKLLAETSSIDDLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOCKLAVGSMHNATRFGRMPQAEKEKLLAETSSIDDLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHITPLOE 238
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGDEKIKFKHITPLOE 300
QY 239 OSKEVAIRIFOGCCFRSVEAVOEITEYAKSIPGFVNLDLNOVTLTKGVHEIITMLAS 298
Db 301 OSKEVAIRIFOGCCFRSVEAVOEITEYAKNIPGFVNLNDLNOVTLTKGVHEIITMLAS 360
QY 299 LMNKDGVLSIEGCGFMTEFLSKLRPGDMEPKFEFAVKNALDELDDSLAIFIAVII 358
Db 361 LMNKDGVLSIEGCGFMTEFLSKLRPGDMEPKFEFAVKNALDELDDSLAIFIAVII 420
QY 359 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLQKMTDLRQIYTEHVOL 418
Db 421 LSGDRPGLLNKPRIEDIDNLQALELQKLNHPESSQLFAKLQKMTDLRQIYTEHVOL 480
QY 419 LQVIKKTETDMSLHPLLOEIKDLY 443
Db 481 LQVIKKTETDMSLHPLLOEIKDLY 505

```

```

RESULT 7
PPAT_CRIGR
ID PPAT_CRIGR STANDARD; PRT; 475 AA.
AC P57797;

```

```

DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARC OR NR1C3.
OS Citreellus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetulus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Liver;
RX MEDLINE=96032361; PubMed=7557447;
  Apoelo C., Pognonec P., Saladin R., Auwerx J., Boulukos K.E.;
  "cDNA cloning and characterization of the transcriptional activities
  of the hamster peroxisome proliferator-activated receptor hppar
  gamma.";
  RL Gene 162:297-302(1995).
CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z30972; CAA83219.1; -.
DR HSSP: P37231; 1FM9.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001723; Stchrnm_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STRDIDFINGER.
DR PRODOM: PD000035; Znf_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; Znf_C4; 1.
DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; zinc-finger; Multigene family; Phosphorylation.
FT DNA_BIND 109 173 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 109 129 C4-TYPE.
FT ZN_FING 146 168 C4-TYPE.
FT DOMAIN 288 475 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 82 82 PHOSPHORYLATION (BY MAPK) (BY
  SIMILARITY)
SO SEQUENCE 475 AA; 54472 MW; BBDCA0704F837ADB CRC64;
Query Match 58.2%; Score 2043; DB 1; Length 475;
Best Local Similarity 84.8%; Pred. No. 3.7e-106;
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;
OY 31 MVDTEPMPNPNFGISSVDSYMEDHSHFDIKPTTVDFSSISPHYEDIDPFTTDPV 90
  |||||
DB 1 MVDTEPMPNPNFGISSVDSYMEDHSHFDIKPTTVDFSSISPHYEDIDPFTTADPV 60
OY 91 ADVKXDLKLOEQSAIKVEPASPYYSEKTOLYN----- 124
  |||||
DB 61 ADVKXDLKLOEQSAIKVEPASPYYSEKTOLYNRPHEPSPNSLMAITRCVCGKASGPH 120

```

```

OY 125 -----RNKCYCRFOKCLAVGMSHNAIRF 148
  |||||
DB 121 YGVNAECGCKGFFRRTRIRKLTYDRCDLNCRIHKKSRKQCRCFOKCLAVGMSHNAIRF 180
OY 149 GMPAPAEKEKLLAEISSDIDQNPESADLRALAKLYOSYIKSPFLTRAKRAAILTGATT 208
  |||||
DB 181 GMPAPAEKEKLLAEISSDIDQNPESADLRALAKLYOSYIKSPFLTRAKRAAILTGATT 240
OY 209 DKSPFVIYDMNSLMGDEKIKFKHITPLQEOSKEVAIRIFQCGCPRSYEAQVEITEYVKS 268
  |||||
DB 241 DKSPFVIYDMNSLMGDEKIKFKHITPLQEOSKEVAIRIFQCGCPRSYEAQVEITEYVKN 300
OY 269 IPGFVNLNDNDVQVTLKYGVEIIVTMLASLNNKQGVLIISGQGFMTREPLKSLRPPGD 328
  |||||
DB 301 IPGFVNLNDNDVQVTLKYGVEIIVTMLASLNNKQGVLIISGQGFMTREPLKSLRPPGD 360
OY 329 FMEPRFEFAVKFNALIEDSDLAIFAVIILSGDRPGLLNKPIEDIDNLLQALELQK 388
  |||||
DB 361 FMEPRFEFAVKFNALIEDSDLAIFAVIILSGDRPGLLNKPIEDIDNLLQALELQK 420
OY 389 LNHPESSOLFALQKMTDLROIVTEHVOLLQVYKTEFDMSLHPLQEIYKDL 443
  |||||
DB 421 LNHPESSOLFALQKMTDLROIVTEHVOLLQVYKTEFDMSLHPLQEIYKDL 475
RESULT 8
PSTAT_RABIT
ID PSTAT_RABIT STANDARD: PRT: 475 AA.
AC 019052;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PPARC OR NR1C3.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=New Zealand white;
RX MEDLINE=97418978; PubMed=9275054;
  Michael L.F., Lazar M.A., Mendelson C.R.;
  "Peroxisome proliferator-activated receptor gamma1 expression is
  induced during cyclic adenosine monophosphate-stimulated
  differentiation of alveolar type II pneumocytes.";
  RL Endocrinology 138:3695-3703(1997).
CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U84893; AAB96380.1; -.
DR HSSP: P37231; 1FM9.
DR InterPro: IPR000536; Hormone_rec.1lg.
DR InterPro: IPR001723; Stchrnm_receptor.
DR InterPro: IPR001628; Znf_C4steroid.
DR Pfam: PF00104; hormone_rec.1.
DR Pfam: PF00105; zf-C4; 1.

```

```

DR PRINTS; PR00398; STRDHOMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; ZNF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Phosphorylation.
FT DNA_BIND 109 173 NUCLEAR_RECEPTOR-TYPE.
FT ZN_FING 109 128 C4-TYPE.
FT ZN_FING 146 168 C4-TYPE.
FT DOMAIN 288 475 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 82 82 PHOSPHORYLATION (BY MAPK) (BY
SO SEQUENCE 475 AA; 54364 MW; FA26829531009E7 CRC64;

Query Match 58.2%; Score 2041; DB 1; Length 475;
Best Local Similarity 84.8%; Pred. No. 4,8e-106;
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;

QY 31 MVDTEPMPPTNFGISSVDSLVSMDHSHSPDIKPFITVDSSISTPHYEDIPFTTDPVY 90
DB 1 MVDTEPMPPTNFGISSVDSLVSMDHSHSPDIKPFITVDSSISAPHYEDLPARADPMV 60
QY 91 ADYKYDLKLOEYOSAIKVEPASPYYSEKTYLN----- 124
DB 61 ADYKYDLKLOEYOSAIKVEPASPYYSEKTYLNKTHEEPSNSLMAIECRVCDKASGFH 120
QY 125 -----RNKCQYRFPCKCLAVGSHNAIRF 148
DB 121 YGVHAGECGKGFRRTRIKLIYRCDLNCRIHKSSRNKCQYCRFOKCLAVGSHNAIRF 180
QY 149 GRMPOAEKEKLEIASSIDOLNPESADLRALKHLYDSYIKSFPLTKAKARAILGKTT 208
DB 181 GRMPOAEKEKLEIASSIDOLNPESADLRALKHLYDSYIKSFPLTKAKARAILGKTT 240
QY 209 DKSPFIYDMSLMDGEDRIKFKHITPLOSKEVAIRIFOGCQFVSVAVOETFEYAKS 268
DB 241 DKSPFIYDMSLMDGEDRIKFKHITPLOSKEVAIRIFOGCQFVSVAVOETFEYAKN 300
QY 269 IPGFVLDLNDQVTLKYGVHEITITMLASLMMKDCVLLISEGCGFMTREFLSLRKPEGD 328
DB 301 IPGFVLDLNDQVTLKYGVHEITITMLASLMMKDCVLLISEGCGFMTREFLSLRKPEGD 360
QY 329 FMPEKEFAVFNALBLDSDLAIFAVIILISGDRGLNVKPIEDIONLQALELOLK 388
DB 361 FMPEKEFAVFNALBLDSDLAIFAVIILISGDRGLNVKPIEDIONLQALELOLK 420
QY 389 LNHPESSQLFAKLLQKMTDLROIIVTEHVOLLQYIKKTEFDMSLHPLLQETKYDLY 443
DB 421 LNHPESSQLFAKLLQKMTDLROIIVTEHVOLLQYIKKTEFDMSLHPLLQETKYDLY 475

RESULT 9
PPT_XENLA STANDARD; PRT; 477 AA.
ID PPT_XENLA STANDARD; PRT; 477 AA.
AC P37234;
DT 01-OCT-1994 (rel. 30; Last sequence update)
DT 01-OCT-1994 (rel. 30; Last sequence update)
DT 16-OCT-2001 (rel. 40; Last annotation update)
DE Peroxisome proliferator activated receptor gamma (PPAR-gamma).
GN PARG OR NP1C3.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodidae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92191267; PubMed=1312391;
RA Dreyer C., Krey G., Keller H., Givel F., Helftenbein G., Wahl W.;
RT "Control of the peroxisomal beta-oxidation pathway by a novel family
of nuclear hormone receptors."

```

```

RL Cell 68:879-887(1992).
RN [2]
RP CHARACTERIZATION.
RX MEDLINE=94100165; PubMed=8274443;
RA Krey G., Keller H., Mahfoudi A., Medin J., Ozato K., Dreyer C.,
RA Wahl W.;
RT "Xenopus peroxisome proliferator activated receptors: genomic
RT organization, response element recognition, heterodimer formation
RT with retinoid X receptor and activation by fatty acids."
RL J. Steroid Biochem. Mol. Biol. 47:65-73(1993).
CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS. KEY REGULATOR
CC OF ADIPOCYTE DIFFERENTIATION AND GLUCOSE HOMEOSTASIS.
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: EXPRESSED MAINLY IN ADIPOSE TISSUE AND KIDNEY.
CC -1- DEVELOPMENTAL STAGE: ADULT.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M84163; AAA49937.1; -.
DR PIR; CA2214; CA2214.
DR HSSP; P37231; 1PM9.
DR TRANSFAC; T01354; -.
DR InterPro; IPR000536; Hormone_rec_1lg.
DR InterPro; IPR001723; Strdhm_receptor.
DR InterPro; IPR001628; Znf_C4steroid.
DR Pfam; PF00104; hormone_rec; 1.
DR Pfam; PF00105; Zf-C4; 1.
DR PRINTS; PR00398; STRDHOMONER.
DR PRINTS; PR00047; STROIDFINGER.
DR PRODOM; PD000035; Znf_C4steroid; 1.
DR SMART; SM00430; HOL1; 1.
DR SMART; SM00399; ZNF_C4; 1.
DR PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Phosphorylation.
FT DNA_BIND 113 177 NUCLEAR_RECEPTOR-TYPE.
FT ZN_FING 113 133 C4-TYPE.
FT ZN_FING 150 172 C4-TYPE.
FT DOMAIN 293 477 LIGAND-BINDING (POTENTIAL).
FT MOD_RES 87 87 PHOSPHORYLATION (BY MAPK) (BY
SO SEQUENCE 477 AA; 54055 MW; 160F87A401C87246 CRC64;

Query Match 41.9%; Score 1471; DB 1; Length 477;
Best Local Similarity 62.7%; Pred. No. 1.5e-74;
Matches 307; Conservative 44; Mismatches 49; Indels 90; Gaps 9;

QY 31 MVDTEPMPPT-NFGISSVDSLVSMDHSHSPDIKPFITVDSSISTPHYEDIP-----FT 84
DB 1 MVDTEPMTSNLNFQNSMDKALDHCOPYDKIRKFTTVDSSINS-HYDOLDEKTFLC 59
QY 85 RTDVPVADKYKDLKLOEYOSAIKVEPASPYYSEKTYQ----- 121
DB 60 RNDQSPIDYKYDLKLOEYOSAIKVEPASPYYSEKTYQ----- 119
QY 122 -----LY-----RNKCQYRFPCKCLAVGSH 143
DB 120 ASGFHYGVHAGECGKGFRRTRIKLIYRCDLNCRIHKSSRNKCQYCRFOKCLAVGSH 179
QY 144 NAIRGRMPOAEKEKLEIASSIDOLNPESADLRALKHLYDSYIKSFPLTKAKARAIL 203

```

```

Db 180 NAFRFRMPAEKEKLAETSSDIDLPESADQRYLAKHLSDYVSKFPLTKAKAGHP 239
Cc 204 TGT--TDSPEFYIMNSLMGEDIKFKHITPLOB-----SKVAIRIFOGCOF 253
Cc 240 DQSHRONSGYRHRHLDADDGGSDGAVRE--PRAQGGGSDNLPALVALR----- 290
Cc 254 RSEVAOETEVAKSIPGFVNLNDQVTLTKYGVHEIITMLASLMMKDGVLISGOGF 313
Cc 291 ---GVRTEIEFAKNIPGFVSLDNLQVTLTKYGVHEIITMLASLMMKDGVLISGOGF 347
Cc 314 MTRFELKSLKPPGDMPEKFEPAVKFNALEDDSLAIFIAVITISGDRPGLLNKPIE 373
Cc 348 MTRFELKSLKPPGDMPEKFEPAVKFNALEDDSLAIFIAVITISGDRPGLLNKPIE 407
Cc 374 DIDNLLQALELQKLNHPSSQLPAKLLQKMTDLRQIVTEHVALQVTKETEDSLHP 433
Cc 408 DIDSLQALELQKLNHPSSQLPAKLLQKMTDLRQIVTEHVALQVTKETEDSLHP 467
Cc 434 LLOEYKDLV 443
Cc 468 LLOEYKDLV 477

RESULT 10
PPAS_HUMAN STANDARD; PRT; 441 AA.
ID 003181;
AC 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peroxisome proliferator activated receptor beta (PPAR-beta)
DE (PPAR-delta) (Nuclear hormone receptor 1) (NUC1) (NUC1).
GN PPARB OR NR1C2 OR PPARD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiinae; Homo.
OX NCBI_TaxID=9606;
RN 1;
RP SEQUENCE FROM N.A.
RX MEDLINE=93078797; PubMed=1333051;
RA Schmidt A., Endo N., Rutledge S.J., Vogel R., Shinar D.,
RA Rodan G.A.;
RT "Identification of a new member of the steroid hormone receptor
RT superfamily that is activated by a peroxisome proliferator and fatty
RT acids."
RL Mol. Endocrinol. 6:1634-1641(1992).
RN 2;
RP SEQUENCE FROM N.A.
RX MEDLINE=20311491; PubMed=10851270;
RA Skogsberg J., Kannisto K., Roshani L., Gagne E., Hamsten A.,
RA Larsson C., Ehrenborg E.;
RT "Characterization of the human peroxisome proliferator activated
RT receptor delta gene and its expression."
RL Int. J. Mol. Med. 6:73-81(2000).
RN 3;
RP SEQUENCE FROM N.A.
RA Phillips S.;
RT Submitted (Aug-1998) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.
CC -!- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: UBIQUITOUS WITH MAXIMAL LEVELS IN PLACENTA AND
CC SKELETAL MUSCLE.
CC -!- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

```

```

Cc the European Bioinformatics Institute. There are no restrictions on its
Cc use by non-profit institutions as long as its content is in no way
Cc modified and this statement is not removed. Usage by and for commercial
Cc entities requires a license agreement (See http://www.isb-sib.ch/announce/
Cc or send an email to license@isb-sib.ch).
Cc -----
Cc EMBL; L07592; AAA36469.1; -
Cc EMBL; AF246303; AAF62553.1; -
Cc EMBL; AF246299; AAF62553.1; JOINED.
Cc EMBL; AF246300; AAF62553.1; JOINED.
Cc EMBL; AF246301; AAF62553.1; JOINED.
Cc EMBL; AF246302; AAF62553.1; JOINED.
Cc EMBL; AL022721; CAB38629.1; -
Cc PIR; A45360; A45360.
Cc HSP; P03372; IHC0.
Cc TRANSFAC; T02745; -
Cc Genew; HGNC; 9235; PPARD.
Cc MIM; 600409; -
Cc InterPro; IPR000536; Hormone_rec_11g.
Cc InterPro; IPR001723; Stdrhm_receptor.
Cc InterPro; IPR001628; Znf_C4steroid.
Cc Pfam; PF00104; hormone_rec.1.
Cc Pfam; PF00105; zf-C4; 1.
Cc PRINTS; PR00398; STRDHOMNER.
Cc PRODOM; PD000035; Znf_C4steroid.1.
Cc SMART; SM00430; HOL1.1.
Cc SMART; SM00399; ZNF_C4; 1.
Cc PROSITE; PS00031; NUCLEAR_RECEPTOR; 1.
Cc Receptor; Transcription regulation; Activator; DNA-binding;
Cc Nuclear protein; Zinc-finger; Multigene family.
Cc FT DNA_BIND 74 138 NUCLEAR RECEPTOR-TYPE.
Cc FT ZN_FING 74 94 C4-TYPE.
Cc FT ZN_FING 111 133 C4-TYPE.
Cc FT DOMAIN 254 441 LIGAND-BINDING (BY SIMILARITY).
Cc SQ SEQUENCE 441 AA; 49903 MW; 94FBB2A4B46521E8 CRC64;

Query Match 31.4%; Score 1101.5; DB 1; Length 441;
Best Local Similarity 63.8%; Pred. No. 3.5e-54;
Matches 208; Conservative 56; Mismatches 61; Indels 1; Gaps 1;

Cc 119 KTYLYRNKCYQCFQKCLAVGSHNAIRGMPQAEKELLAET-SSDIDLPESADL 177
Cc 116 KTKKRNKCYQCFQKCLAVGSHNAIRGMPQAEKELLAET-SSDIDLPESADL 175
Cc 178 RALAKLYDYSIKSPPLTAKARAILTGKTDKSPVYIDMSLMGCEKIKFKHITPQ 237
Cc 176 KAFSKHYNAVYLNKFNWTKKARSILTGKASTAPRVIHDIETLWQAEGLWKKLVNGL 235
Cc 238 EQSKEVAIRIFOGCQFSEVAOETEVAKSIPGFVNLNDQVTLTKYGVHEIITMLA 297
Cc 236 PYKEISVHVFRCQCTYETVELTEFPAKSIPFSFLNDQVTLTKYGVHEIIFAMLA 295
Cc 298 SLMMKGVGLISGQGFMTREFLKSLKPPGDMPEKFEPAVKFNALEDDSLAIFIAV 357
Cc 296 SYVKNKGLVANGSGVTRFEFLKSLKPPGDMPEKFEPAVKFNALEDDSLAIFIAV 355
Cc 358 ILSGDRPGLLNKPIEDIDNLLQALELQKLNHPSSQLPAKLLQKMTDLRQIVTEHQ 417
Cc 356 ILGDRPGLLNKPIEDIDNLLQALELQKLNHPSSQLPAKLLQKMTDLRQIVTEHQ 415
Cc 418 LLOVIRKTTEDMSLHPDLOEYKDLV 443
Cc 416 MMQRIKKTETETSLHPLDLOEYKDMY 441

RESULT 11
PPAR_RAT STANDARD; PRT; 468 AA.
ID P37230;
AC 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

```

```

DE Peroxisome proliferator activated receptor alpha (PPAR-alpha).
GN PPARA OR NR1C1 OR PPAR.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92262498; PubMed=1316614;
RA Goettlicher M., Midmark E., Li Q., Gustafsson J.-A.;
RT "Fatty acids activate a chimera of the clofibric acid-activated
RT receptor and the glucocorticoid receptor."
RL Proc. Natl. Acad. Sci. U.S.A. 89:4653-4657(1992).
RN [2]
RP PHOSPHORYLATION.
RX MEDLINE=96426218; PubMed=8828512;
RA Shalev A., Slegrist-Kaiser C.A., Yen P.M., Wahl W., Burger A.G.,
RA Chin W.W., Meier C.A.;
RT "The peroxisome proliferator-activated receptor alpha is a
RT phosphoprotein: regulation by insulin."
RL Endocrinology 137:4499-4502(1996).
CC -FUNCTION: RECEPTOR THAT BINDS PEROXISOME PROLIFERATORS SUCH AS
CC HYPOLIPEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACTL-COA
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE
CC PEROXISOXOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.
CC -SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.
CC -TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN LIVER AND KIDNEY.
CC -PTM: PHOSPHORYLATED.
CC -SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.
CC NRI SUBFAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isdb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC CC
DR EMBL: M88592; AAA41918.1; -.
DR PIR: A45288; A45288.
DR HSSP: P03372; IHGQ.
DR TRANSFAC: T00091; -.
DR InterPro: IPR000536; Hormone_rec_1fg.
DR InterPro: IPR001723; Sterhmrn_receptor.
DR InterPro: IPR001628; Znfc_C4steroid.
DR Pfam: PF00104; hormone_rec_1.
DR Pfam: PF00105; zf-C4; 1.
DR PRINTS: PR00398; STRDHOMONER.
DR PRINTS: PR00047; STROIDFINGER.
DR ProDom: PD000035; Znfc_C4steroid; 1.
DR SMART: SM00430; HOL1; 1.
DR SMART: SM00399; ZnFc_C4; 1.
DR PROSITE: PS00031; NOCLEAR_RECEPTOR; 1.
KW Receptor; Transcription regulation; Activator; DNA-binding;
KW Nuclear protein; Zinc-finger; Multigene family; Phosphorylation.
FT DNA_BIND 102 166 NUCLEAR RECEPTOR-TYPE.
FT ZN_FING 102 122 C4-TYPE.
FT FT 139 161 C4-TYPE.
FT ZN_FING 139 161 C4-TYPE.
FT DOMAIN 281 468 LIGAND-BINDING (POTENTIAL).
SQ SEQUENCE 468 AA; 52377 MW; 2A89E7D715C8DBA9 CRC64;
Query Match 31.1%; Score 1091.5; DB 1; Length 468;
Best Local Similarity 49.2%; Pred. No. 1,4e-53;
Matches 237; Conservative 70; Mismatches 112; Indels 63; Gaps 10;
OY 7 DSPIDP----SDSPFTDLANISOGLMTVMVTLEMPFWPTNGCISVSVDVSYMEDSHSFDI 62
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 5 ESPICPLPSLEADDELSPLESEEFLOENG-----NIQETISQSLGEFSGSGTSF 51
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
63 kPF-----TTVFSSISTPHYEDIPF--FRTD-----PVAD 92

```

Db	52	ADYQYLGSCPGSEGSVITDITLSPASSPSSVSCPAVPSTDES	PGMALNIECRICGDKASG	1111
Oy	93	KYDYLKLOE-----YQSAIKYEPASPPYYSE-----	KTOLYNRNKCOCYCRPQCLAYGMS	142
Db	112	YHYGVAHECGCKGFFRRIRIRLKLKLA-----YD	CDSCKTQKRNKRKCOCYCRPHKCLSGMS	167
Oy	143	HNARFEGMPOAEKEKELIAEI-SSDIDQLNPESADLR	ALAKHLYDSYIKSEPLTKARARA	201
Db	168	HNARFEGMPPSEKAKLAEILITCEHDLKDESETADL	SKARIHEATLKFNEMNKVKARV	227
Oy	202	ILTGRTTQKSPFYIYDNMSLMMGEKIKFKHITPLQ	EDSKREVARIRIQGOCFSEVAEVOE	266
Db	228	ILAGTSTNNPPVHIDMETLMAEETLVAKNVANCV	-NKEAEVRFPFCQCMSEVYTE	286
Oy	262	ITEYAKSIPGCVNIDLNOQVLTLLKGVHEIITYIT	MLASLMNKDGLVISEGCGFMFREFLKS	321
Db	287	LTFEPKALPGFANDLNDQVLTLLKGYEALFTML	SLMKNKGMLIYAGNCFITREFLKN	346
Oy	322	LRRPGEFMEERKEFEFAVKFNALIEDDSDLAIF	AVIILLSGDRPLNVKPIEDIQDMLQ	381
Db	347	LRRPCCDIMEKPEDFAMKFNLEIDDSDISLF	VAATICCGDRPLNIGYTEKQEGIVH	406
Oy	362	ALELDLKLNHSESSOLFKKLLQKNTDLRQYTE	HVOLLQYTKETDMSLHPLLOEYKD	441
Db	407	VKLHLQSNHDDPTFLFKLLQKMWDLQLTVE	HAOLVQVIKKTESDAALHPILQETIRD	466
Oy	442	LY 443		
Db	467	MY 468		
RESULT 12				
PPAR_MOUSE				
ID	PPAR_MOUSE	STANDARD:	PRT:	468 AA.
AC	P23204;			
DT	01-NOV-1991 (Rel. 20, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Peroxisome proliferator activated receptor alpha (PPAR-alpha).			
GN	PPARA OR NR1C1 OR PPAR.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxId=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=91015382; PubMed=2129546;			
RT	Isseemann I., Green S.;			
RT	"Activation of a member of the steroid hormone receptor superfamily			
RT	by peroxisome proliferators.";			
RL	Nature 347:645-650(1990).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv;			
RX	MEDLINE=94168583; PubMed=8123021;			
RT	Gearing K.L., Crickmore A., Gustafsson J.-A.;			
RT	"Structure of the mouse peroxisome proliferator activated receptor			
RT	alpha gene.";			
RL	Biochem. Biophys. Res. Commun. 199:255-263(1994).			
RN	[3]			
RP	SEQUENCE OF 413-468 FROM N.A.			
RC	STRAIN=Swiss Webster; TISSUE=Liver;			
RX	MEDLINE=96061953; PubMed=7586749;			
RA	Jones P.S., Savory R., Barratt P., Bell A.R., Gray T.J.B.,			
RA	Jeukens N.A., Gilbert D.J., Copeland N.G., Bell D.R.;			
RT	"Chromosomal localisation, inducibility, tissue-specific expression			
RT	and strain differences in three murine			
RT	peroxisome proliferator-activated-receptor genes.";			
RL	Eur. J. Biochem. 233:219-226(1995).			
CC	-1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND, THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA			



OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
 PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
 CC SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN LIVER, KIDNEY AND  
 CC HEART. VERY WEAKLY EXPRESSED IN BRAIN AND TESTIS.  
 CC -1- DEVELOPMENTAL STAGE: IT APPEARS FIRST AT DAY 13.5 POSTCONCEPTION,  
 CC AND INCREASES UNTIL BIRTH.  
 CC -1- DISEASE: PEROXISOME PROLIFERATORS ARE A DIVERSE GROUP OF  
 CC CHEMICALS THAT INCLUDE HYPOLIPIDEMIC DRUGS, HERBICIDES AND  
 CC INDUSTRIAL PLASTICISERS. ADMINISTRATION OF THESE CHEMICALS TO  
 CC ROBERTS RESULTS IN THE DRAMATIC PROLIFERATION OF HEPATIC  
 CC PEROXISOMES AS WELL AS LIVER HYPERTPLASIA.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
 CC NRI SUBFAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL, X57638; CAA40856.1; -;  
 DR EMBL, X75289; CAA53042.1; -;  
 DR EMBL, X75290; CAA53042.1; JOINED.  
 DR EMBL, X75291; CAA53042.1; JOINED.  
 DR EMBL, X75292; CAA53042.1; JOINED.  
 DR EMBL, X75293; CAA53042.1; JOINED.  
 DR EMBL, X75294; CAA53042.1; JOINED.  
 DR EMBL, X89577; CAA61754.1; -;  
 DR PIR, S11659; S11659.  
 DR HSSP, P03372; IHCO.  
 DR TRANSFAC: T00694; -;  
 DR MGD, MGI:104740; Ppara.  
 DR InterPro: IPR000536; Hormone\_rec\_Lig.  
 DR InterPro: IPR001723; Steroid\_receptor.  
 DR InterPro: IPR001628; znf\_C4steroid.  
 DR Pfam: PF00104; hormone\_rec\_1.  
 DR Pfam: PF00105; zf-C4\_1.  
 DR PRINTS: PR00398; STRDHOMONER.  
 DR PRINTS: PR00047; STROIDFINGER.  
 DR ProDom: PD000035; znf\_C4steroid\_1.  
 DR SMART: SM00430; HOL1; 1.  
 DR SMART: SM00399; znf\_C4\_1.  
 DR PROSITE: PS00031; NUCLEAR\_RECEPTOR\_1.  
 DR Receptor: Transcription regulation; Activator; DNA-binding;  
 KM Nuclear protein; Zinc-finger; Multigene family.  
 KM DNA\_BIND 102 166 NUCLEAR RECEPTOR-TYPE.  
 FT ZN\_FING 102 122 C4-TYPE.  
 FT ZN\_FING 139 161 C4-TYPE.  
 FT DOMAIN 281 468 LIGAND-BINDING (POTENTIAL).  
 FT CONFLICT 75 75 A->R (IN REF.1).  
 SQ SEQUENCE 468 AA; 52347 MW; 2950A5191C610B6B CRC64;

Query Match 30.9%; Score 1083.5; DB 1; Length 468;  
 Best Local Similarity 48.4%; Pred. No. 3-8e-53;  
 Matches 236; Conservative 65; Mismatches 112; Incls 75; Gaps 9;

OY 7 DSPIDP-----ESDFTDITANISODMTWDTMPMPPTNFISSVSDYSVMEDHSIEDI 62  
 Db 5 ESDIPPLPLEADDESPLEEFLOEMG-----NIOEISOSIGESSSGFG 51  
 OY 63 KFE-----TYDPSISPHYEDIPRTDPVADVADKLQYOSAIKYE 109  
 Db 52 ADYQLGSCPGSEGSYTDITLSPASSPSVSCPVI---PASTD-----ESPGSALNTE 101  
 OY 110 -----PASPPY-----SEKTOLYNNKQCYGRFOKC 136  
 Db 102 CRIGCDKASGYHYGVHACGCGKGFRRITRLKLYVDKCRSKICKKNNKQCYGRFHNC 161  
 OY 137 LAVGSHNAIRGRMPQAEKELAEI-SSDIDQINPESADLRALAKLYDSYIKSFPLT 195

Db 162 LSVGSHNAIRGRMPRSEKAKLKAELLCEHDLNDETFADLSGKRHEHYLNFMNN 221  
 OY 196 KAKARILITGKTDDSKSPFIYDMNSLMGEDIKFKHTIPLOQSKVAIRIFQCCOPRS 255  
 Db 222 KVKARVILAGKTSNNPPYIHMETLCNAEKTLYAKMVAANGVE-DKEAEVRFHCCQCKS 280  
 OY 256 VEAVQITEYAKSIPGFVNLNDQVTLTKYGVHIIYTMLASLNKDKGVLSIEGQEMT 315  
 Db 281 VETVELTEFAAIGFANLDNDQVTLTKYVEAIFMTLSSLNKKGMLIANGFIT 340  
 OY 316 REPLKSLKRPEDMEPKPEFAVKFNALFLDSDLAIFLAVIILSGRPLNVKPIEDI 375  
 Db 341 REFLNLRKPFCDIMEPFDFAKMFNALELDDSDISLHVAALICCGDPGLNIGYIKL 400  
 OY 376 QNLLQALELQKLNHPRESSQFAKLQKMDLRQIVLHENVOLLOVIRKTTENDSLHPL 435  
 Db 401 QEGIVHVLKHLQSNHPDQTFLEPKLQKMDLRQIVLHENVOLVIRKTTESDALHPL 460  
 OY 436 QEYKDYLY 443  
 Db 461 QEYKDYLY 468

RESULT 13  
 PPAS\_MOUSE STANDARD: PRT; 440 AA.  
 ID PPAS\_MOUSE P37239;  
 AC P35396; P37239;  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 39, Last annotation update)  
 DE 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Peroxisome proliferator activated receptor beta (PPAR-beta)  
 DE (PPAR-delta) (Nuclear hormone receptor 1) (NUC1).  
 GN PPARB OR NR1C2 OR PPAR.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Adipocyte;  
 RX MEDLINE=95138211; PubMed=7836471;  
 RA Amri E.-Z., Bonhio F., Allaud G., Abumrad N.A., Grimaldi P.A.;  
 RT "Cloning of a protein that mediates transcriptional effects of fatty  
 RT acids in preadipocytes. Homology to peroxisome proliferator-activated  
 RT receptors.";  
 RL J. Biol. Chem. 270:2367-2371(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=94316694; PubMed=8041794;  
 RA Klierer S.A., Forman B.M., Blumberg B., Ong E.S., Borgmeyer U.,  
 RT Mangelsdorf D.J., Unesono K., Evans R.M.;  
 RT "Differential expression and activation of a family of murine  
 RT peroxisome proliferator-activated receptors.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 91:7355-7359(1994).  
 RN [3]  
 RP SEQUENCE OF 1-145 FROM N.A.  
 RC STRAIN=BALB/c; TISSUE=Brain;  
 RX MEDLINE=94059089; PubMed=8240342;  
 RA Chen F., Law S.W., O'Malley B.W.;  
 RT "Identification of two mPPAR related receptors and evidence for the  
 RT existence of five subfamily members.";  
 RL Biochem. Biophys. Res. Commun. 196:671-677(1993).  
 CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS  
 CC HYPOLIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
 CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
 CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
 CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
 CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: HEART, ADRENAL AND INTESTINE.  
 CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.



```

CC -----
CC      NRI SUBFAMILY.
CC -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (See http://www.isb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC -----
CC      EMBL: L28116; AAA63394.1; -.
CC      EMBL: U10375; AAA19972.1; -.
CC      EMBL: U01665; AAA03332.1; ALT_INIT.
CC      HSP: P03372; IHCO.
CC      TRANSFAC: T04781; -.
CC      MGD: MGI:101884; Ppard.
CC      InterPro: IPR000536; Hormone_rec_1lg.
CC      InterPro: IPR001723; Steroid_receptor.
CC      Pfam: PF00104; hormone_rec; 1.
CC      Pfam: PF00105; zf-C4; 1.
CC      PRINTS: PR00398; STROHORMONER.
CC      PRINTS: PR00047; STROIDPFINGER.
CC      PRODOM: PD000035; Znf_C4steroid; 1.
CC      SMART: SM00430; HOL1; 1.
CC      SMART: SM00399; Znf_C4; 1.
CC      PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.
CC      Receptor: Transcription regulation; Activator; DNA-binding;
CC      Nucleic protein; Zinc-finger; Multigene family.
CC      DNA_BIND 73 137 NUCLEAR RECEPTOR-TYPE.
CC      ZN_FING 73 93 C4-TYPE.
CC      ZN_FING 110 132 C4-TYPE.
CC      DOMAIN 253 440 LIGAND-BINDING (POTENTIAL).
CC      CONFID 149 150 EA -> DG (IN REF. 2).
CC      SEQUENCE 440 AA; 49715 MW; 58E0F595DD0193DA CRC64;

Query Match      30.98; Score 1082.5; DB 1; Length 440;
Best Local Similarity 62.6%; Pred. No. 3.9e-53;
Matches 204; Conservative 59; Mismatches 62; Indels 1; Gaps 1;

QY 119 KTLQNLNKKQCYGRFOKCLAVGSMHAIIRGRMPQAEKELAEI-SSDIDQINPEASADL 177
DB 115 KICKKNNKQCYGRFOKCLAVGSMHAIIRGRMPQAEKELAEI-SSDIDQINPEASADL 174
QY 178 RALAKHLHYDSYKSPFLTKAKARAIITGKTTDSPVIYDMSIMMGEDKIRKHTIPQD 237
DB 175 KAKSKHYNAYLNFMNMTKKARSILGKSSHNAPEYIHDIETLMOAEKGLVAKQVLNGL 234
QY 238 ESKREVAIRIFGCGQFRSVAQVEITEYAKSIPGFYNLDINDQVTLTKYGVHEIITYTMA 297
DB 235 PYNNEISVHFYRQSGTVEYRELTEFAKNIPNFSLSFLNDQVTLTKYGVHEIIFAMLA 294
QY 298 SLNNKGVLSIEGCGFMTRFELSLKRPDGMPEKPEFVKNALDELDSOLAIFITAVI 357
DB 295 STYNNKGLVANSQGVTEHFEELSLKRPESDIEKPEFVKNALDELDSOLAIFITAVI 354
QY 358 ILSGDRPGLINVPIDIDQNLQALELQNLNHPSSOLFAKLQMTDRLQIYTEHVO 417
DB 355 ILLGDRPGLINVPIDIDQNLQALELQNLNHPSSOLFAKLQMTDRLQIYTEHVO 414
QY 418 LLDQVYKTEFDSLHPLDQIYDLY 443
DB 415 MMQMLKTESEETLHPLDQIYDLY 440

```

```

RESULT 14
PPAR_HUMAN      STANDARD;          PRT;      468 AA.
AC  Q07869; Q92689; Q16241; Q92486; Q9Y3N1;
DT  01-OCT-1994 (Rel. 30, Created)
DT  15-JUL-1998 (Rel. 36, Last sequence update)
DT  15-JUN-2002 (Rel. 41, Last annotation update)
DE  Peroxisome proliferator activated receptor alpha (PPAR-alpha).

```

```

GN  PPARA OR NR1C1 OR PPAR.
OS  Homo sapiens (Human).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX  NCBI_TaxID=9606;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RX  MEDLINE=93277839; PubMed=7684926;
RT  "Shen T., Yi H.F., McBride O.W., Gonzales F.J.;
RT  "CDNA cloning, chromosomal mapping, and functional characterization
RT  of the human peroxisome proliferator activated receptor.";
RL  Biochemistry 32:5598-5604(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RL  Roberts R.A., James N.H., Woodvatt N.J., Macdonald N., Tugwood J.D.;
RN  Submitted (Aug-1996) to the EMBL/Genbank/DBD databases.
RN  [3]
RP  SEQUENCE FROM N.A.
RC  TISSUE=Liver;
RX  MEDLINE=95071923; PubMed=7981125;
RT  Mukherjee R., Jow L., Noonan D., McDonnell D.P.;
RT  "Human and rat peroxisome proliferator activated receptors (PPARs)
RT  demonstrate similar tissue distribution but different responsiveness
RT  to PPAR activators.";
RN  J. Steroid Biochem. Mol. Biol. 51:157-166(1994).
RN  [4]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=20057165; PubMed=10591208;
RN  Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RN  Clamp M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.R.,
RN  Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasly O.P.,
RN  Bird C.P., Blakey S.E., Bridgman A.M., Buck D., Burgess J.,
RN  Burrill W.D., Burton C., Carter N.P., Chen Y., Clark G.,
RN  Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RN  Conroy D., Cordy N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RN  Dhami P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RN  Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RN  Gilbert J.G.R., Goward M.E., Grafham D.V., Griffiths M.N.D., Hall C.,
RN  Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RN  Hunt S.E., Jones M.C., Kersey J., Kimberley A.M., King A.,
RN  Lalit G.K., Langford C.F., Leverisa M.A., Lloyd C., Lloyd D.M.,
RN  Marilyn I.D., Mashreghi-Mohammadi M., Matthews L.H., Mccann O.T.,
RN  Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RN  Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillimore B.J.C.T.,
RN  Phillips S.H., Plumb R.W., Ramsey H., Ramsey Y., Rogers L., Ross M.T.,
RN  Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RN  Soderlund C., Spragon L., Steward C.A., Sulston J.E., Swann R.M.,
RN  Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RN  Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RN  Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RN  Minoshima A., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RN  Shintani A., Shibuya K., Yoshitake Y., Aoki N., Mitsuyama S.,
RN  Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RN  Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RN  Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.,
RN  Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaull S., Sloan D., Song L.,
RN  Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RN  Zhan M., Zhang G., Chissoe S., Murray J., Miller N., Minx P.,
RN  Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RN  Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RN  Hinds K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RN  Scheet P., Walker C., Wamsley A., Woldmann P., Pepin K., Neilson J.,
RN  Korf I., Bedell J.A., Hillier L., Mardis E., Waterston B., Wilson R.,
RN  Emanuel B.S., Shaikh T., Kurahashi H., Saita S., Budarf M.L.,
RN  McDermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RN  Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyrard M., Kedra D.,
RN  Seroussi E., Franssen I., Tapia I., Brydner C.E., O'Brien K.P.,
RN  Wilkerson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RN  Tilahun Y., Wright H.;
RT  "The DNA sequence of human chromosome 22.";
RT  Nature 402:489-495(1999).
CC  -i- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS

```

```
CC HYPO-LIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: SKELETAL MUSCLE, LIVER, HEART AND KIDNEY.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC NRI SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: L02932; AAA36468.1; -  
CC EMBL: Y07619; CA68898.1; -  
CC DR EMBL: S74349; AAB32649.1; -  
CC DR EMBL: AL049856; CAB42862.2; -  
CC DR EMBL: AL078611; CAB44427.1; -  
CC DR PIR: A42889; A49289.  
CC DR HSSP: P03372; IHCO.  
CC DR TRANSFAC: T02726; -  
CC DR Genev: HGNC:9232; PPARA.  
CC DR MIM: 170998; -  
CC DR InterPro: IPR000536; Hormone_rec_1lg.  
CC DR InterPro: IPR001723; Stdhnm_receptor.  
CC DR InterPro: IPR001628; Znf_C4steroid.  
CC DR Pfam: PF00104; hormone_rec; 1.  
CC DR Pfam: PF00105; zf-C4; 1.  
CC DR PRINTS: PR00398; STRDHOMONER.  
CC DR PRINTS: PR00047; STROIDFINGER.  
CC DR ProDom: PD000035; Znf_C4steroid; 1.  
CC DR SMART: SM00430; HOL1; 1.  
CC DR SMART: SM00399; Znf_C4; 1.  
CC DR PROSITE: PS00031; NUCLEAR_RECEPTOR; 1.  
CC DR Receptor: Transcription regulation: Activator: DNA-binding;  
CC KW Nuclear protein; Zinc-finger; Multigene family.  
CC KM DNA_BIND 102 166 NUCLEAR RECEPTOR-TYPE.  
CC FT ZN_FING 102 122 C4-TYPE.  
CC FT ZN_FING 139 161 C4-TYPE.  
CC FT DOMAIN 281 468 LIGAND-BINDING (POTENTIAL).  
CC FT CONFLICT 71 71 T -> M (IN REF. 2).  
CC FT CONFLICT 123 123 K -> M (IN REF. 2).  
CC FT CONFLICT 268 268 A -> V (IN REF. 1).  
CC FT CONFLICT 296 296 G -> A (IN REF. 1).  
CC FT CONFLICT 444 444 V -> A (IN REF. 2).  
CC SQ SEQUENCE 468 AA; 52225 MW; 850846FD51ADA883 CRC64;  
  
Query Match 30.7%; Score 1076.5; DB 1; Length 468;  
Best Local Similarity 49.8%; Pred. No. 9; 2e-53;  
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;  
  
OY 31 MVTPEMPFMTN-----FGISSYDLSVMEDSHSHDKPFTY---68  
DB 1 MVTPEFLCLSLPAGDLESPISEFLMGNTIOETISOSIGESSOSFGTEFOYIGSC 60  
OY 69 ---DFSSIS-TPHYEDLPRTDPVADYKYDKLOEYOS-AIKVE-----PASPPYV-116  
DB 61 PGSDGSVITPTLSPASSPSVTPYVPG-----SVDESPGALNIECRIGDKASGVHYG 115  
OY 117 -----SEKTOLYRNKCQYCRFOKCLAVGMSHNAIRFGR 150  
DB 116 VHACEGCKGFFRRRTIRKLIVYDKDSCKIOKRNKNCQYCRFHCKCLVGMSHNAIRFGR 175  
OY 151 MPAQEKELIAEI-SSDIQNPESADIRALAHLDYSIKSPFLRKARALILGTGTTD 209  
DB 176 MPSESEKALAEILTCHEIDSEITADLKSLARIYEAYLKNPNMKVKARVILASDN 235  
OY 210 KSPFVIYDNMSLMMGEDKIKFKHITFLOESKEVAIRIFGCGFRSVEAVEQITEVAKSI 269
```

```
DB 236 NPPVHIDMETLCAEKTIVAK-LVANGIONKKEVAIRIFHCOCSTSEVITELTEFFAKAI 294  
OY 270 PGEVNDLNDQVTLTKYGVHEIITYTLASLMNKDGLVISEGCGFMTRREPLKSLRKPGDF 329  
DB 295 PGFANLDLNDQVTLTKYGVHEIITYTLASLMNKDGLVISEGCGFMTRREPLKSLRKPGDF 354  
OY 330 MEPEFPAVKFNLEDDSPSLAFIVIIISGDRPGLNPKPIEDIODNLQALELQKL 389  
DB 355 MEPEFPAVKFNLEDDSPSLAFIVIIISGDRPGLNPKPIEDIODNLQALELQKL 414  
OY 390 NHPESSQLFAKLQKMTDLRQIYTEHVQLQVIRKTEETDMSLPLQEIYKDY 443  
DB 415 NHPDPIFLPKLLQKNMADLQVTEHMQVLQIIRKTESDAALPHLQEIYRDWY 468  
  
RESULT 15  
PPAR_XENLA  
ID PPAR_XENLA STANDARD; PRT; 474 AA.  
AC P37232;  
DT 01-OCT-1994 (Rel. 30, Last Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Peroxisome proliferator activated receptor alpha (PPAR-alpha).  
GN PPARA OR NR1C1.  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae;  
OX Xenopodinae; Xenopus.  
NCBI_TaxID=6355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92191267; PubMed=1312391;  
RA Dreyer C., Krey G., Keller H., Givel F., Helltenbein G., Wahl W.;  
RT "Control of the peroxisomal beta-oxidation pathway by a novel family  
RL of nuclear hormone receptors.";  
RN Cell 68:879-887 (1992).  
RP CHARACTERIZATION.  
RX MEDLINE=94100165; PubMed=8274443;  
RA Krey G., Keller H., Mahfoudi A., Medin J., Ozato K., Dreyer C.,  
RA Wahl W.;  
RT "Xenopus peroxisome proliferator activated receptors: genomic  
RT organization, response element recognition, heterodimer formation  
RT with retinoid x receptor and activation by fatty acids.";  
RL J. Steroid Biochem. Mol. Biol. 47:65-73 (1993).  
CC -1- FUNCTION: RECEPTOR THAT BIND PEROXISOME PROLIFERATORS SUCH AS  
CC HYPO-LIPIDEMIC DRUGS AND FATTY ACIDS. ONCE ACTIVATED BY A LIGAND,  
CC THE RECEPTOR BINDS TO A PROMOTER ELEMENT IN THE GENE FOR ACYL-COA  
CC OXIDASE AND ACTIVATES ITS TRANSCRIPTION. IT THEREFORE CONTROLS THE  
CC PEROXISOMAL BETA-OXIDATION PATHWAY OF FATTY ACIDS.  
CC -1- SUBUNIT: HETERODIMER WITH THE RETINOID X RECEPTOR.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.  
CC -1- DEVELOPMENTAL STAGE: OCYTES, EMBRYOS, AND ADULTS.  
CC -1- SIMILARITY: BELONGS TO THE NUCLEAR HORMONE RECEPTOR FAMILY.  
CC NRI SUBFAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL: M84161; AAA9935.1; -  
CC PIR: A42214; A42214.  
CC HSSP: P03372; IHCO.  
CC DR TRANSFAC: T01352; -  
CC DR InterPro: IPR000536; Hormone_rec_1lg.  
CC DR InterPro: IPR001723; Stdhnm_receptor.  
CC DR InterPro: IPR001628; Znf_C4steroid.
```

DR	Pfam: PF00104, hormone_rec. 1.
DR	Pfam: PF00105, zf-C4. 1.
DR	PRINTS: PR00398, STRDHOMONER.
DR	PRINTS: PR00047, STROIDINGER.
DR	Prodom: PD000035, znf_Casteroid. 1.
DR	SMART: SM00430; HOL1. 1.
DR	SMART: SM00399; znf_C4. 1.
DR	PROSITE: PS00031, NUCLEAR_RECEPTOR. 1
KW	Receptor; transcription regulation; Activator; DNA-binding;
KW	Nuclear protein; zinc-finger; Multigene family.
FT	DN_BIND 109 173 NUCLEAR RECEPTOR-TYPE.
FT	ZN_FING 109 129 C4-TYPE.
FT	ZN_FING 146 168 C4-TYPE.
FT	DOMAIN 287 474 LIGAND-BINDING (POTENTIAL).
SO	SEQUENCE 474 AA; 5286 MW; 4DB94131BD43483 CRC64;

Query Match	30.2%	Score 1060.5	DB 1	length 474;
Best Local Similarity	63.58%	Pred. No. 7.1e-52;		
Matches 207, Conservative	51;	Mismatches 65;	Indels 3;	Gaps 3;

QY	119	KTOLYNNRKCQCYCFQKCLAVGMSHNAIRGQMAQAKKELLWEI-SSDDIOLNPEESADL	177
Db	151	KIOKNNRKCQCYCFREKCLAVGMSHNAIRGQMPRESEKATLAEVLMLCDDQDVWDSQWADL	210
QY	178	RALKLHLVDSTIKSFPLTKAKARILNGTKTTKDSPEFYIYOMNSLMGEGDKIKKRNHTPLQ	237
Db	211	LSLARLLYDALIKNFNMNKKVAKARILNGKSN--PPFYIHMETLQMAEKLTVAK-LVANG	268
QY	238	EOSKEVALIRIPOGOCFERSVEAVCITEETAKSIPGEVMLDINDQVTLKYGVNHETIYMLA	297
Db	269	IQNNEAEVRIFHCQCSVSEVETVELTEFASIPGTEFLDINDQVTLKYGVETEMFEMLA	328
QY	298	SLMKDGVLLISEGGGFMTRFLFKSLKRPFGDPEMEKTEFAVKNALDSDLAIFTAIVT	357
Db	329	SVNMKDGVLVYVGGFTREFLKLRLRPIDDMMPKEEFAMKFNALDSDLSLFFAAL	388
QY	358	ILSDRGELNVKRPIDEDIODMLLOALLOELKLNPESSQAFKILQKMTLRLRIVTEHYO	417
Db	389	ICGDDRGVLNVPISIEKMQESIYAVLKLHQSNNPDSFLFPRKLQKMAIDLRLQVTEHQAQ	448
QY	418	LLOYIKTEEDMSLHPLLOELKYDLY	443
Db	449	LVOTIKTEEDALHPLLOELRYDMY	474

Search completed: February 25, 2003, 03:59:00  
Job time : 23 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 25, 2003, 03:56:01 ; Search time 38 seconds  
(without alignments)  
1740.539 Million cell updates/sec

Title: US-09-931-007A-1  
Perfect score: 3508  
Sequence: 1 MGFTLGDSPIDPESDSFTDT.....KTETDMSLHPLQLQETKDLX 688

Scoring table: BIOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	505	JC4859	peroxisome prolif
2	2160.5	61.6	504	JF0280	peroxisome prolif
3	2153	61.4	505	A54101	peroxisome prolif
4	2132	60.8	505	JC5777	peroxisome prolif
5	2043	58.2	475	JC4264	peroxisome prolif
6	2041	58.2	475	JF0279	peroxisome prolif
7	1471	41.9	477	C42214	peroxisome prolif
8	1100.5	31.4	441	A45360	steroid hormone re
9	1091.5	31.1	468	A45288	peroxisome prolif
10	1083.5	30.9	468	JC2085	peroxisome prolif
11	1082.5	30.9	440	I55442	peroxisome prolif
12	1076.5	30.7	440	JC4530	peroxisome prolif
13	1076.5	30.7	468	I56603	peroxisome prolif
14	1074.5	30.6	468	A49289	peroxisome prolif
15	1060.5	30.2	474	A42214	peroxisome prolif
16	946.5	27.0	396	B42214	peroxisome prolif
17	659	18.8	141	PC4290	peroxisome prolif
18	408	11.6	576	A57048	nuclear receptor R
19	397	11.3	578	S52913	nuclear receptor R
20	392	11.2	1394	B34598	ecdysone-induced p
21	391	11.1	579	A57057	orphan nuclear hor
22	385.5	11.0	1237	A34598	ecdysone-induced p
23	384.5	11.0	508	A30226	thyroid/steroid ho
24	379	10.8	614	A32608	thyroid hormone re
25	375.5	10.7	483	I55219	steroid hormone re
26	371.5	10.6	1443	S05979	steroid hormone re
27	354.5	10.1	711	S43464	ecdysone-induced p
28	351	10.0	685	C56591	E75 B steroid rece
29	341.5	9.7	468	A56856	retinoid-related o

30	341.5	9.7	523	2	A53196	orphan hormone nuc
31	341.5	9.7	548	2	C53196	orphan hormone nuc
32	341.5	9.7	556	2	B53196	orphan hormone nuc
33	338.5	9.6	523	2	S68517	nuclear hormone re
34	331	9.4	448	2	B56558	retinoic acid rece
35	330	9.4	464	2	A56558	retinoic acid rece
36	328	9.4	459	2	A41977	retinoic acid rece
37	327.5	9.3	452	2	S78481	retinoic acid rece
38	327.5	9.3	458	2	S06123	retinoic acid rece
39	326.5	9.3	448	2	A43786	retinoic acid rece
40	326.5	9.3	455	2	S13512	retinoic acid rece
41	325.5	9.3	560	2	JC2494	nuclear receptor R
42	324.5	9.3	448	2	S02827	retinoic acid rece
43	323.5	9.2	444	2	I51256	retinoic acid rece
44	323.5	9.2	445	2	A56043	steroid hormone re
45	321.5	9.2	955	4	C40045	probable transcript

ALIGNMENTS

RESUT 1  
JC4859  
peroxisome proliferator-activated receptor gamma-2 - human  
C:Species: Homo sapiens (man)  
C:Date: 15-Aug-1996 #sequence\_revision 18-Oct-1996 #text\_change 20-Sep-1999  
C:Accession: JC4859; S71351  
R:Elbrecht, A.; Chen, Y.; Cullinan, C.A.; Hayes, N.; Leibowitz, M.D.; Moller, D.E.; B  
Biochem. Biophys. Res. Commun. 224, 431-437, 1996  
A>Title: Molecular cloning, expression and characterization of human peroxisome prol  
A:Reference number: JC4859; MUID:96295505; PMID:8702406  
A:Contents: fat  
A:Accession: JC4859  
A:Molecule type: mRNA  
A:Residues: 1-505 <ELB>  
A:CROSS-references: GB:063415; NID:g1432176; PIDN:AAB04028.1; PID:g1432177  
R:Rambe, K.G.; Tugwood, J.D.  
Eur. J. Biochem. 239, 1-7, 1996  
A>Title: A human peroxisome-proliferator-activated receptor-gamma is activated by ind  
A:Reference number: S71351; MUID:96305559; PMID:8706692  
A:Accession: S71351  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 29-505 <LAMB>  
A:CROSS-references: EMBL:X90563; NID:g1480099; PIDN:CA62152.1; PID:e258302; PID:g149  
A:Experimental source: liver  
C:Comment: This receptor plays a role in adipocyte differentiation and in regulating  
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
C:Keywords: DNA binding; heterodimer; receptor; zinc finger  
F:1-505/Product: peroxisome proliferator-activated receptor gamma-2 #status predicted  
F:29-505/Product: peroxisome proliferator-activated receptor gamma-1 #status predicted  
F:137-423/Domain: erba transforming protein homology <RBA>

Query Match 63.6%; Score 2231; DB 2; Length 505;  
Best Local Similarity 87.7%; Pred. No. 9.3e-117;  
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY	1	MGFTLGDSPIDPESDSFTDTLSANISQEXTWVTEMPPTNFGISSVDLSWMEHSHSF	60
DB	1	MGFTLGDSPIDPESDSFTDTLSANISQEXTWVTEMPPTNFGISSVDLSWMEHSHSF	60
QY	61	DKPPTVDFSSISPHYEDIPFTPTDPVADYKXDKLQEOSAIKVPASPYSEKT	120
DB	61	DKPPTVDFSSISPHYEDIPFTPTDPVADYKXDKLQEOSAIKVPASPYSEKT	120
QY	121	QLYN-----	124
DB	121	QLYN-----	124
QY	125	-----RKKCQCRQKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQINPESADLR	178
DB	181	RIHKSRNKCQCRQKCLAVGSMHNAIRFGMPQAEKEKLLAEISSDIDQINPESADLR	240

```

OY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 238
    |||||||
DB 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 300
OY 239 OSKEVAIRIFGCGFRSVAVOETITEYAKSIRGFVNLNDQYTLTKYGVHEIITYMLAS 298
    |||||||
DB 301 OSKEVAIRIFGCGFRSVAVOETITEYAKSIRGFVNLNDQYTLTKYGVHEIITYMLAS 360
OY 299 LMKKDVLLISEGCGFMTREFLKSIRKPFQDFMEPKFEFAVKFNALDSDLAIFIAVII 358
    |||||||
DB 361 LMKKDVLLISEGCGFMTREFLKSIRKPFQDFMEPKFEFAVKFNALDSDLAIFIAVII 420
OY 359 LSGDRPGLLVKPRFIEDIONLLOALELQKLKHPRESSQLFAKLQKMTDLROIIVTEHVOL 418
    |||||||
DB 421 LSGDRPGLLVKPRFIEDIONLLOALELQKLKHPRESSQLFAKLQKMTDLROIIVTEHVOL 480
OY 419 LOYIKKETDMSLHPLLOETIKDLY 443
    |||||||
DB 481 LOYIKKETDMSLHPLLOETIKDLY 505

```

## RESULT 2

```

JE0280
peroxisome proliferator-activated receptor gamma2 - pig
C:Species: Sus scrofa domestica (domestic pig)
C:Date: 05-Feb-1999 #sequence_revision 05-Feb-1999 #text_change 21-Jan-2000
C:Accession: JE0280
R:Grindflek, E.; Sundvold, H.; Klungland, H.; Lien, S.
Biochem. Biophys. Res. Commun. 249, 713-718, 1998
A:Title: Characterization of porcine peroxisome proliferator-activated receptors gamma1
A:Reference number: JE0279; MUID:98401156; PMID:9731203
A:Accession: JE0280
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-504 <GRD>
A:Cross-references: GB:AJ006756
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
E:136-422/Domain: erba transforming protein homology <ERB>

```

```

Query Match          61.6%; Score 2160.5; DB 2; Length 504;
Best Local Similarity 85.1%; Pred. No. 7.7e-113;
Matches 430; Conservative 7; Mismatches 5; Indels 63; Gaps 2;

OY 1 MGETLGDSPIDPSDSFTDILSANISOEMTWDTMPMPNFGISSVDLSVMDHSHSF 60
    |||||||
DB 1 MGETLGDSPIDPSDSFTDILSANISOEMTWDTMPMPNFGISSVDLSVMDHSHSF 59
OY 61 DIKPFPTVDFSSISPHYEDIPRTDPVADYKYDKLQEOSAKIKVPASPPYSEKT 120
    |||||||
DB 60 DIKPFPTVDFSSISPHYEDIPRTDPVADYKYDKLQEOSAKIKVPASPPYSEKT 119
OY 121 QLYN----- 124
    |||
DB 120 QLYNKRHEEPSNLMAIEGRVCGDKASGFHYGVHACEGCKGFFRRIRIKLIYDRCDLNC 179
OY 125 -----RNCQYCRROCKLAVGKSHNAIRFGMPQAEKKLLAETSSDIDQLPESADIR 178
    |||||||
DB 180 RIHKRSRNCQYCRFOCKLAVGKSHNAIRFGMPQAEKKLLAETSSDIDQLPESADIR 239
OY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 238
    |||||||
DB 240 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMGEDIKIKFHHITPLOE 299
OY 239 OSKEVAIRIFGCGFRSVAVOETITEYAKSIRGFVNLNDQYTLTKYGVHEIITYMLAS 298
    |||||||
DB 300 OSKEVAIRIFGCGFRSVAVOETITEYAKSIRGFVNLNDQYTLTKYGVHEIITYMLAS 359
OY 299 LMKKDVLLISEGCGFMTREFLKSIRKPFQDFMEPKFEFAVKFNALDSDLAIFIAVII 358
    |||||||
DB 360 LMKKDVLLISEGCGFMTREFLKSIRKPFQDFMEPKFEFAVKFNALDSDLAIFIAVII 419
OY 359 LSGDRPGLLVKPRFIEDIONLLOALELQKLKHPRESSQLFAKLQKMTDLROIIVTEHVOL 418
    |||||||

```

```

DB 420 LSGDRPGLLVKPRFIEDIONLLOALELQKLKHPRESSQLFAKLQKMTDLROIIVTEHVOL 479
OY 419 LOYIKKETDMSLHPLLOETIKDLY 443
    |||||||
DB 480 LOYIKKETDMSLHPLLOETIKDLY 504

```

## RESULT 3

```

A54101
peroxisome proliferator-activated receptor gamma - mouse
C:Species: Mus musculus (house mouse)
C:Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 20-Sep-1999
C:Accession: A54101; A57740; A49294; JN0881; S53748
R:Tontono, P.; Hu, E.; Graves, R.A.; Budavari, A.I.; Spiegelman, B.M.
Genes Dev. 8, 1224-1234, 1994
A:Title: mPARGamma2: tissue-specific regulator of an adipocyte enhancer.
A:Reference number: A54101; MUID:95011536; PMID:7926726
A:Accession: A54101
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-505 <TON>
A:Cross-references: EMBL:U09138; NID:9500639; PIDN:AAA62277.1; PID:9500640
A:Experimental source: adipose tissue
R:Kliwer, S.A.; Forman, B.M.; Blumberg, B.; Ong, E.S.; Borgmeyer, U.; Mangelsdorf, D
Proc. Natl. Acad. Sci. U.S.A. 91, 7355-7359, 1994
A:Title: Differential expression and activation of a family of murine peroxisome prol
A:Reference number: A57740; MUID:94316694; PMID:8041794
A:Accession: A57740
A:Molecule type: mRNA
A:Residues: 31-382, 'S', 384-505 <KLI>
A:Cross-references: GB:U10374; NID:9507776; PIDN:AAA19971.1; PID:9514307; GB:U10375
A:Experimental source: liver
R:Zhu, Y.; Alvares, K.; Huang, Q.; Rao, M.S.; Reddy, J.K.
J. Biol. Chem. 268, 26817-26820, 1993
A:Title: Cloning of a new member of the peroxisome proliferator-activated receptor ge
A:Reference number: A49294; MUID:94086482; PMID:8262913
A:Accession: A49294
A:Molecule type: mRNA
A:Residues: 31-505 <ZHU>
A:Cross-references: GB:U01841; NID:9454000; PIDN:AA052134.1; PID:9454001
A:Experimental source: liver
R:Chen, F.; Law, S.W.; O'Malley, B.W.
Biochem. Biophys. Res. Commun. 196, 671-677, 1993
A:Title: Identification of two mPAP related receptors and evidence for the existence
A:Reference number: JN0881; MUID:94059089; PMID:8240342
A:Accession: JN0881
A:Molecule type: mRNA
A:Residues: 31-212, 'DR', 215-280, 'S', 282, 'F', 284-382, 'S', 384-441, 447-496, 'F', 498-505 <
A:Experimental source: heart
R:Tontono, P.; Graves, R.A.; Budavari, A.I.; Erdjument-Bromage, H.; Lui, M.; Hu, E.;
Nucleic Acids Res. 22, 5628-5634, 1994
A:Title: Adipocyte-specific transcription factor ARF6 is a heterodimeric complex of t
A:Reference number: S53748; MUID:95140625; PMID:7838715
A:Accession: S53748
A:Molecule type: Protein
A:Residues: 66-85, 146-160 <TOM>
A:Experimental source: adipose tissue
C:Complex: heterodimer; peroxisome proliferator-activated receptor gamma and retinoid
C:Function:
A:Description: transcription regulation of adipogenesis; activates transcription in r
A>Note: binding of the heterodimer to its recognition element induces DNA bending
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
C:Keywords: DNA binding; heterodimer; nucleus; receptor; transcription regulation; z1
F:137-423/Domain: erba transforming protein homology <ERBA>
F:139-159/Region: zinc finger CCCC motif
F:160-224/Region: DNA binding #status predicted
F:176-198/Region: zinc finger CCCC motif

```

```

Query Match          61.4%; Score 2153; DB 2; Length 505;
Best Local Similarity 84.0%; Pred. No. 2e-112;
Matches 424; Conservative 9; Mismatches 10; Indels 62; Gaps 1;

OY 1 MGETLGDSPIDPSDSFTDILSANISOEMTWDTMPMPNFGISSVDLSVMDHSHSF 60

```

```

Db      1 MGELTIGDSEVPDEHGAFAADALPMSTSQELTIVDTMEPTNPIFGISSVDLSMEDHSHSF 60
        |||
QY      61 DIKPTFTVDFESSISTPHNEDIPFTRTDPVADYKYKDLKLOEYOSAIKVEPASPYYSEKT 120
        |||
Db      61 DIKPTFTVDFESSISAPHNEDIPFTRADPVADYKYDKLQKEYOSAIKVEPASPYYSEKT 120
        |||
QY      121 QLYN----- 124
        |||
Db      121 QLYNRPHEEPSNLSMAIECRVCGDKASGFHYGVNAHECGCKGFRRTRILKLIYRCDLNC 180
        |||
QY      125 -----RNNCOYCRFOKCLAVGSMHNAIRFGMRPAEKELAEISSDIDQINPSSADLR 178
        |||
Db      181 RIHKRSRNKCYCRFOKCLAVGSMHNAIRFGMRPAEKELAEISSDIDQINPSSADLR 240
        |||
QY      179 ALAKHLUDYSIKSFPLTKAKARAILTGKTTDQSPVYIDMSLMNGEDKIKFKHTLPLOE 238
        |||
Db      241 ALAKHLUDYSIKSFPLTKAKARAILTGKTTDQSPVYIDMSLMNGEDKIKFKHTLPLOE 300
        |||
QY      239 OSKEVAIRIFOCQFRSVEAVOEITEYAKSIPGFVNLDLNDQVTLTKXGVHEIITYMLAS 298
        |||
Db      301 OSKEVAIRIFOCQFRSVEAVOEITEYAKNIPGFINLNDLNDQVTLTKXGVHEIITYMLAS 360
        |||
QY      299 LANKGVGLISEQGWTREFLKSLKPPGDMPEKPFERFAVKFNALIEDDSLAIFITAVII 358
        |||
Db      361 LANKGVGLISEQGMTRFEFLKLNKPPGDMPEKPFERFAVKFNALIELDDSDLAIFIAVII 420
        |||
QY      359 LSGDPRGLINVPRIEDIDONLQALELOLKLNHPPSSQLFAKLKMTDLROIYEHVOL 418
        |||
Db      421 LSGDPRGLINVPRIEDIDONLQALELOLKLNHPPSSQLFAKLKMTDLROIYEHVOL 480
        |||
QY      419 LQVIRKTTDMSLHPLLOEIKYKLY 443
        |||
Db      481 LQVIRKTTDMSLHPLLOEIKYKLY 505
        |||

RESULT 4
JC5777
peroxisome proliferator-activated receptor gamma 2 - bovine
C.Species: Bos primigenius taurus (cattle)
C.Date: 24-Jan-1998 #sequence_revision 13-Mar-1998 #text_change 24-Sep-1999
C.Accession: JC5777
A:Accession: JC5777
A.Molecule type: mRNA
A.Residues: 1-505 <SUN>
A.Cross-references: G6:Y12420; NID:q2653405; PIDN:CAA73033.1; PID:ej188692; PID:q2653400
A.Note: The sequences of residues 301-320 and 321-340 are interchanged in the authors' t
C.Comment: By ligand activation, this protein regulates the expression of many genes enc
hondrial beta-oxidation, and adipocyte differentiation.
C.Genetics:
A:Map position: 22
C:Superfamily: unassigned erba-related proteins; erba transforming protein homology
F.1-505/P:Product: peroxisome proliferator-activated receptor gamma 2 #status predicted <
F.31-505/P:Product: peroxisome proliferator-activated receptor gamma 1 #status predicted <
F.137-422/Domain: erba transforming protein homology <RBA>

Query Match 60.8%; Score 2132; DB 2; Length 505;
Best Local Similarity 83.6%; Pred. No. 2.9e-11;
Matches 422; Conservative 11; Mismatches 10; Indels 62; Gaps 1;

QY      1 MGELTIGDSEVPDEHGAFAADALPMSTSQELTIVDTMEPTNPIFGISSVDLSMEDHSHSF 60
        |||
Db      1 MGELTIGDALIDESSEPAFTVARSARTSOEITWDTMPMPNFGISSVDLSMDHSHAF 60
        |||
QY      61 DIKPTFTVDFESSISPHNEDIPFTRTDPVADYKYDKLQEOYSAIKVEPASPYYSEKT 120
        |||
Db      61 DIKPTFTVDFESSISPHNEDIPFPADPVADYKYDKLQEOYSAIKVEPASPYYSEKT 120
        |||
QY      121 QLYN----- 123
        |||

```

QY	124	-----NRNRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	178
QY	124	181 RHKKSRNKCQYCRQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	240
QY	179	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	238
QY	241	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	300
QY	239	QSEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	298
QY	301	PSKEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	360
QY	299	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	358
QY	361	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	420
QY	359	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	418
QY	421	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	480
QY	419	LOVYIKTETDMSLHPLOEIKYKDL 443	
QY	481	LOVYIKTETDMSLHPLOEIKYKDL 505	
QY	121	QLYSKPHEBPSLMAIECRVCGDKASGEHYGVACBEGCKGEFRRRTIRLKLITYRCDLNC	180
QY	124	-----NRNRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	178
QY	181	RHKKSRNKCQYCRQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	240
QY	179	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	238
QY	241	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	300
QY	239	QSEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	298
QY	301	PSKEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	360
QY	299	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	358
QY	361	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	420
QY	359	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	418
QY	421	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	480
QY	419	LOVYIKTETDMSLHPLOEIKYKDL 443	
QY	481	LOVYIKTETDMSLHPLOEIKYKDL 505	
QY	121	QLYSKPHEBPSLMAIECRVCGDKASGEHYGVACBEGCKGEFRRRTIRLKLITYRCDLNC	180
QY	124	-----NRNRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	178
QY	181	RHKKSRNKCQYCRQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	240
QY	179	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	238
QY	241	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	300
QY	239	QSEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	298
QY	301	PSKEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	360
QY	299	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	358
QY	361	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	420
QY	359	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	418
QY	421	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	480
QY	419	LOVYIKTETDMSLHPLOEIKYKDL 443	
QY	481	LOVYIKTETDMSLHPLOEIKYKDL 505	
QY	121	QLYSKPHEBPSLMAIECRVCGDKASGEHYGVACBEGCKGEFRRRTIRLKLITYRCDLNC	180
QY	124	-----NRNRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	178
QY	181	RHKKSRNKCQYCRQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	240
QY	179	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	238
QY	241	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	300
QY	239	QSEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	298
QY	301	PSKEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	360
QY	299	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	358
QY	361	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	420
QY	359	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	418
QY	421	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	480
QY	419	LOVYIKTETDMSLHPLOEIKYKDL 443	
QY	481	LOVYIKTETDMSLHPLOEIKYKDL 505	
QY	121	QLYSKPHEBPSLMAIECRVCGDKASGEHYGVACBEGCKGEFRRRTIRLKLITYRCDLNC	180
QY	124	-----NRNRCQYCFQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	178
QY	181	RHKKSRNKCQYCRQKCLAVGMSHNAIRFGMPQAEKEKLAETSSDIDLPESADLR	240
QY	179	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	238
QY	241	ALAKHLVDSYIKSPFLYAKARAILTGKTTDKSPFVIYDMSLMGDEKIKFKHITPLQEE	300
QY	239	QSEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	298
QY	301	PSKEVAIRIFQGCQFRSVEAEOETETAKSIPGEFNLNDQVYTLKGVHEIITYTLAS	360
QY	299	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	358
QY	361	LNKNGGVLSIEGCGMTREFELSKRPGDFMEPEKFEFAVENALIEDDSQALFIAVII	420
QY	359	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	418
QY	421	LSGDRPGLLVNRPIDIEDIDNLLQALELQKLNPRESSOLFALLQKMTDLQOYTEHVQL	480
QY	419	LOVYIKTETDMSLHPLOEIKYKDL 443	
QY	481	LOVYIKTETDMSLHPLOEIKYKDL 505	
QY	121	QLYSKP	

QY 329 FMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIEDIQNLQALQELQK 388  
 |||||||  
 Db 361 FMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIEDIQNLQALQELQK 420  
 |||||||  
 QY 389 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHPLQELQYDLY 443  
 |||||||  
 Db 421 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHPLQELQYDLY 475  
 |||||||

RESULT 6  
 JF0279  
 Peroxisome proliferator-activated receptor gamma 1 - pig  
 C:Species: Sus scrofa domestica (domestic pig)  
 C:Date: 05-Feb-1999 #sequence\_revision 05-Feb-1999 #text\_change 21-Jan-2000  
 C:Accession: JF0279  
 R:Grindflek, E.; Sundvold, H.; Klungland, H.; Lien, S.  
 Biochem. Biophys. Res. Commun. 249, 713-718, 1998  
 A:Title: Characterization of porcine peroxisome proliferator-activated receptors gamma1  
 A:Reference number: JF0279; MUID:98401156; PMID:9731203  
 A:Accession: JF0279  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-475 <GR>  
 A:Cross-references: GB:A006758  
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
 F:107-393/Domain: erba transforming protein homology <ERB>

Query Match 58.2%; Score 2041; DB 2; Length 475;  
 Best Local Similarity 85.1%; Pred. No. 3,1e-106;  
 Matches 404; Conservative 5; Mismatches 4; Indels 62; Gaps 1;

QY 31 MVDTEMPFPTNFGISSVDSLVMDHSHFDIKPFTTVDFSSISPTPHYEDIPFTTRDPVY 90  
 |||||||  
 Db 1 MVDTEMPFPTNFGISSVDSLVMDHSHFDIKPFTTVDFSSISPTPHYEDIPFTTRDPVY 60  
 |||||||  
 QY 91 ADYKDYDLQEOYSAIKVEPASPYYSEKTOLYN----- 124  
 |||||||  
 Db 61 ADYKDYDLQEOYSAIKVEPASPYYSEKTOLYNKRHEPSSNLMAIECRVCGDKASGFH 120  
 |||||||  
 QY 125 -----RNKCQYCRQKCLAVGMSHNAIRF 148  
 |||||||  
 Db 121 YGVNAHCEGCKGFERRTRIRKLILYDRCDLNCRIHKSRNKCQYCRQKCLAVGMSHNAIRF 180  
 |||||||  
 QY 149 GRMPQAEKELKLAETSSIDQNLNPESADLRALAKHLVDYIKSPFLTKAKARAILTGKTT 208  
 |||||||  
 Db 181 GRMPQAEKELKLAETSSIDQNLNPESADLRALAKHLVDYIKSPFLTKAKARAILTGKTT 240  
 |||||||  
 QY 209 DKSPFVIYDMSLMMGEDKIKFKHITPLQEOQKEVAIRIFGCGFRSVAVOETETAKS 268  
 |||||||  
 Db 241 DKSPFVIYDMSLMMGEDKIKFKHITPLQEOQKEVAIRIFGCGFRSVAVOETETAKN 300  
 |||||||  
 QY 269 IPGFVNLNDNDVTLTKYGVHEIITYMLASLNNKGVLSIEGQGWTRERFLSKLRPFED 328  
 |||||||  
 Db 301 IPGFVNLNDNDVTLTKYGVHEIITYMLASLNNKGVLSIEGQGWTRERFLSKLRPFED 360  
 |||||||  
 QY 329 FMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIEDIQNLQALQELQK 388  
 |||||||  
 Db 361 FMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIEDIQNLQALQELQK 420  
 |||||||  
 QY 389 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHPLQELQYDLY 443  
 |||||||  
 Db 421 LNHPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHPLQELQYDLY 475  
 |||||||

RESULT 7  
 C42214  
 Peroxisome proliferator-activated receptor gamma chain - African clawed frog  
 N:Alternate names: PPAR-gamma  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999  
 C:Accession: C42214  
 R:Dreyer, C.; Krey, G.; Keller, H.; Givel, F.; Helftenbein, G.; Wahli, W.

Cell 68, 879-887, 1992  
 A:Title: Control of the peroxisomal beta-oxidation pathway by a novel family of nucle  
 A:Reference number: A42214; MUID:92191267; PMID:1312391  
 A:Accession: C42214  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-477 <DR>  
 A:Cross-references: GB:M4163; NID:9214667; PID:AAA49937.1; PID:9214668  
 A>Note: sequence extracted from NCBI backbone (NCBI:88295)  
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
 C:Keywords: DNA binding; transcription regulation; zinc finger  
 F:111-395/Domain: erba transforming protein homology <ERBA>

Query Match 41.9%; Score 1471; DB 2; Length 477;  
 Best Local Similarity 62.7%; Pred. No. 1.4e-74;  
 Matches 307; Conservative 44; Mismatches 49; Indels 90; Gaps 9;

QY 31 MVDTEMPFPTNFGISSVDSLVMDHSHFDIKPFTTVDFSSISPTPHYEDIP-----FT 84  
 |||||||  
 Db 1 MVDTEMPFPTNFGISSVDSLVMDHSHFDIKPFTTVDFSSISPTPHYEDIP-----FT 59  
 |||||||  
 QY 85 RNDPVVADYKDYDLQEOYSAIKVEPASPYYSEKTO----- 121  
 |||||||  
 Db 60 RNDQSPIDYKDYDLQEOYSAIKVEPASPYYSEKTO----- 119  
 |||||||  
 QY 122 -----LY-----NNKCQYCRQKCLAVGMSH 143  
 |||||||  
 Db 120 ASGFHYGVNAHCEGCKGFERRTRIRKLILYDRCDLNCRIHKSRNKCQYCRQKCLAVGMSH 179  
 |||||||  
 QY 144 NAIRFGMPQAEKELKLAETSSIDQNLNPESADLRALAKHLVDYIKSPFLTKAKARAIL 203  
 |||||||  
 Db 180 NAIRFGMPQAEKELKLAETSSIDQNLNPESADLRALAKHLVDYIKSPFLTKAKARAIL 239  
 |||||||  
 QY 204 TPKT---TKSPFVIYDMSLMMGEDKIKFKHITPLQEO-----SKEVAIRIFGCGF 253  
 ||:::|  
 Db 240 DQSHRQNSRGYTRHDLDDGGSDQGAVER-PRAEQGGGDSNLPAISVALR----- 290  
 ||:::|  
 QY 254 RSEVAOETETAKSIPGFVNLNDNDVTLTKYGVHEIITYMLASLNNKGVLSIEGQGF 313  
 ||:::|  
 Db 291 ---GGVRETTETAKNIPGFVNLNDNDVTLTKYGVHEIITYMLASLNNKGVLSIEGQGF 347  
 ||:::|  
 QY 314 MTRERFLSKRPFDFMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIE 373  
 |||||||  
 Db 348 MTRERFLSKRPFDFMEPKFEFAVKNALDELDDSLAIFAVIITLLSGDRPGLNVKPIE 407  
 |||||||  
 QY 374 DIQNDLQALEQLKLNPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHP 433  
 |||||||  
 Db 408 DIQNDLQALEQLKLNPESSQLFAKLQKMTDLROIVTEHVOLLQVYIKKFTETDMSLHP 467  
 |||||||  
 QY 434 LLQETIKDLY 443  
 |||||||  
 Db 468 LLQETIKDLY 477  
 |||||||

RESULT 8  
 A45360  
 steroid hormone receptor delta - human  
 N:Alternate names: nuclear hormone receptor 1 (NUC1); peroxisome proliferator-activat  
 C:Species: Homo sapiens (man)  
 C:Date: 10-Jun-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Oct-1997  
 C:Accession: A45360  
 R:Schmidt, A.; Endo, N.; Rutledge, S.J.; Vogel, R.; Shinar, D.; Rodan, G.A.  
 Mol. Endocrinol. 6, 1634-1641, 1992  
 A:Title: Identification of a new member of the steroid hormone receptor superfamily t  
 A:Reference number: A45360; MUID:93078797; PMID:1333051  
 A:Accession: A45360  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: nucleic acid  
 A:Residues: 1-441 <SCH>  
 A:Experimental source: osteosarcoma SAOS-2/B10 cells  
 A>Note: sequence extracted from NCBI backbone (NCBI:118801)  
 C:Superfamily: unassigned erba-related proteins; erba transforming protein homology  
 C:Keywords: DNA binding; steroid hormone receptor; transcription regulation; zinc fin







Db 295 STVNKDGILLVANGSGFVTHEFLRSIRKPFSDIIEPKFEFAVFNALVELVSDLALEFIAM 354  
QY 358 IISGDRPGLLNKPIEDIDONLLOALELOKLNHPESQOLFAPKLOKMTDLROIVTEHQ 417  
Db 355 IICGDRPGMLNVPQVEAIDDTILOALEFHLQVNHDSQYTLFPKLOKMDLROLVTEHAQ 414  
QY 418 LLOVIRKTEPDMSLHPLLOEIKYDL 443  
Db 415 MMQWLKTESFLLHPLLOEIKYDMY 440

RESULT 13  
156603  
peroxisome proliferator activated receptor alpha - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 20-Sep-1999  
C:Accession: I56603  
R:MuKherjee, R.; Jow, L.; Noonan, D.; McDonnell, D.P.  
J. Steroid Biochem. Mol. Biol. 51, 157-166, 1994  
A:Title: Human and rat peroxisome proliferator activated receptors (PPARs) demonstrate s  
A:Reference number: I56603; MUID:95071923; PMID:7981125  
A:Accession: I56603  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-468 <RES>  
A:Cross-references: GB:S74349; NID:9765239; PIDN:AA832649.1; PID:9765240  
C:Genetics:  
A:Gene: hPPARalpha  
C:Superfamily: unassigned erba-related proteins: erba transforming protein homology  
C:Keywords: zinc finger  
F:100-386/Domain: erba transforming protein homology <ERBA>

Query Match 30.7%; Score 1076.5; DB 2; Length 468;  
Best Local Similarity 49.8%; Pred. No. 1.2e-52;  
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;

QY 31 MVDTEMPMPN-----FGISSVDSLVMEDSHSFDIKPFTTV--- 68  
Db 1 MVDTEPLCLPLSLPGLADLESPLEEFLOEMGNIOEISQIGEDSSGSGFTGYGLSC 60  
QY 69 ---DSSIS-TPHYEDIPTRTDPVADYKYLKQEVOS-AIKVE-----PASPPY- 116  
Db 61 PGSDGSVITDITLSPASSPSVTPVPPG-----SVDESPGALNIECRICGDKASGYHNG 115  
QY 117 -----SEKTOLYNNKCOYCFQFCOCLAVGSMHNAIRGR 150  
Db 116 VIACGCGKFFFRTRIRLKYDKCRSCKIQKNNKCOYCFHKLCLSGVSMHNAIRGR 175  
QY 151 MPOAEKELIAEI-SSDIDQNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTD 209  
Db 176 MPRSEKAKIAELILNCEHIEDSEFADLSLAKRIYEALKNFNNNKVKARVILSGKASN 235  
QY 210 KSPFYIYDMSNLAMGEDKIKFKHITPLOSKEVAIRIFGCGFQSVAVQDITEYAKSI 269  
Db 236 NPPFYIHMETLCMAEKLTVAK-LVANGIONKEAEVRIFHCOCQSVETVETLFEAKAI 294  
QY 270 PGFVNLDLNDQYTLTKYGHETIITMLASIMNKDGLVSEGGFMREFLKLRRPFGDF 329  
Db 295 PGFANLDLNDQYTLTKYGHETIITMLASIMNKDGLVAVGSGFTREFLKLRRPFCII 354  
QY 330 MPRKEFAVFNALVELDSDLAIFIAVITLLSGDRPGLNVKPIEDIDONLLOALELOK 389  
Db 355 MPRKEDFAKFNALVELDSDISLFAVAITICGDRPGLNVGHIEKMGESIVAVLRLHQS 414  
QY 390 NHPESQOLFAPKLOKMTDLROLVTEHVOLLQYIKTEPDMSLHPLLOEIKYDL 443  
Db 415 NHPDIFLFPKLOKMDLROLVTEHAQVQIKTESDALHPLLOEIKYDMY 468

RESULT 14  
A49289  
peroxisome proliferator-activated receptor - human  
C:Species: Homo sapiens (man)

C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 20-Sep-1999  
C:Accession: A49289  
R:Sher, T.; Yi, H.F.; McBride, O.W.; Gonzalez, F.J.  
Biochemistry 32, 5598-5604, 1993  
A:Title: cDNA cloning, chromosomal mapping, and functional characterization of the hu  
A:Reference number: A49289; MUID:93277839; PMID:7684926  
A:Accession: A49289  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-468 <SHC>  
A:Cross-references: GB:L02932; NID:9307340; PIDN:AAA36468.1; PID:9307341  
A>Note: authors translated the codon AAC for residue 33 as Asp  
C:Genetics:  
A:Map position: 22  
C:Superfamily: unassigned erba-related proteins: erba transforming protein homology  
C:Keywords: DNA binding; transcription regulation; zinc finger  
F:100-386/Domain: erba transforming protein homology <ERBA>

Query Match 30.6%; Score 1074.5; DB 2; Length 468;  
Best Local Similarity 49.8%; Pred. No. 1.5e-52;  
Matches 236; Conservative 65; Mismatches 106; Indels 67; Gaps 9;

QY 31 MVDTEMPMPN-----FGISSVDSLVMEDSHSFDIKPFTTV--- 68  
Db 1 MVDTEPLCLPLSLPGLADLESPLEEFLOEMGNIOEISQIGEDSSGSGFTGYGLSC 60  
QY 69 ---DSSIS-TPHYEDIPTRTDPVADYKYLKQEVOS-AIKVE-----PASPPY- 116  
Db 61 PGSDGSVITDITLSPASSPSVTPVPPG-----SVDESPGALNIECRICGDKASGYHNG 115  
QY 117 -----SEKTOLYNNKCOYCFQFCOCLAVGSMHNAIRGR 150  
Db 116 VIACGCGKFFFRTRIRLKYDKCRSCKIQKNNKCOYCFHKLCLSGVSMHNAIRGR 175  
QY 151 MPOAEKELIAEI-SSDIDQNPESADLRALAKHLYDSYIKSFPLTKAKARAILTGKTTD 209  
Db 176 MPRSEKAKIAELILNCEHIEDSEFADLSLAKRIYEALKNFNNNKVKARVILSGKASN 235  
QY 210 KSPFYIYDMSNLAMGEDKIKFKHITPLOSKEVAIRIFGCGFQSVAVQDITEYAKSI 269  
Db 236 NPPFYIHMETLCMAEKLTVAK-LVANGIONKEAEVRIFHCOCQSVETVETLFEAKAI 294  
QY 270 PGFVNLDLNDQYTLTKYGHETIITMLASIMNKDGLVSEGGFMREFLKLRRPFGDF 329  
Db 295 PGFANLDLNDQYTLTKYGHETIITMLASIMNKDGLVAVGSGFTREFLKLRRPFCII 354  
QY 330 MPRKEFAVFNALVELDSDLAIFIAVITLLSGDRPGLNVKPIEDIDONLLOALELOK 389  
Db 355 MPRKEDFAKFNALVELDSDISLFAVAITICGDRPGLNVGHIEKMGESIVAVLRLHQS 414  
QY 390 NHPESQOLFAPKLOKMTDLROLVTEHVOLLQYIKTEPDMSLHPLLOEIKYDL 443  
Db 415 NHPDIFLFPKLOKMDLROLVTEHAQVQIKTESDALHPLLOEIKYDMY 468

RESULT 15  
A42214  
peroxisome proliferator-activated receptor alpha chain - African clawed frog  
N:Alternate names: PPAR-alpha  
C:Species: Xenopus laevis (African clawed frog)  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 20-Sep-1999  
A:Accession: A42214  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-474 <DRS>  
A:Cross-references: GB:M84161; NID:9214663; PIDN:AAA49935.1; PID:9214664  
A>Note: sequence extracted from NCBI backbone (NCBI:88292)  
C:Superfamily: unassigned erba-related proteins: erba transforming protein homology





```
Db 121 QLYNRNKCQYCRFOKCLAVGMSHNAIRFGMPQAEKEKLLAEISSDIDQNPESADLRAL 180
QY 181 AKHLUDYSIKSPFLTKAKARAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQEOS 240
Db 181 AKHLUDYSIKSPFLTKAKARAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQEOS 240
QY 241 KEVAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITMLASLM 300
Db 241 KEVAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITMLASLM 300
QY 301 NKGGVILISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIILS 360
Db 301 NKGGVILISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVIILS 360
QY 361 GDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEHVOLLQ 420
Db 361 GDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEHVOLLQ 420
QY 421 VIKKTETDMSLHPLLOEIRYKDLVAMAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITP 480
Db 421 VIKKTETDMSLHPLLOEIRYKDLVAMAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITP 480
QY 481 LOBOSKVAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYM 540
Db 481 LOBOSKVAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITYM 540
QY 541 LASLMNKDGLVISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFIA 600
Db 541 LASLMNKDGLVISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFIA 600
QY 601 VILSGDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEH 660
Db 601 VILSGDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEH 660
QY 661 VOLLOVYIKKTETDMSLHPLLOEIRYKDLX 688
Db 661 VOLLOVYIKKTETDMSLHPLLOEIRYKDLX 688

RESULT 2
US-09-765-111A-16
; Sequence 16, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PBPARGAMMA NUCLEIC ACID MOLECULES
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-765-111A-16

Query Match 63.6%; Score 2231; DB 10; Length 505;
Best Local Similarity 87.7%; Pred. No. 8.6e-154;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;
```

```
Db 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVVADYKYDKLQEOYSAIKVEBPASPPYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKPHEEPSNLSMAIECVCGDKASGFHYGVHACBGCGFFRRIRIRLKIYDRCDLNC 180
QY 125 -----RNKCOYCRFOKCLAVGMSHNAIRFGMPQAEKEKLLAEISSDIDQNPESADLR 178
Db 181 RHKKSNNKCQYCRFOKCLAVGMSHNAIRFGMPQAEKEKLLAEISSDIDQNPESADLR 240
QY 179 ALAKHLUDYSIKSPFLTKAKARAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQ 238
Db 241 ALAKHLUDYSIKSPFLTKAKARAILTGKTTDKSPFIYDMNSLMGDEKIKRKHITPLQ 300
QY 239 QSKVEAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 298
Db 301 QSKVEAIRIFOGCOPRSVEAVOEITEYAKSIPGFVNLNDQVTLTKYGVHEIITMLAS 360
QY 299 LMKNDGLVISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVII 358
Db 361 LMKNDGLVISEGQGFMTREFLSLRKPFGEFMEPKFEFAVKFNALDELDDSDLAIFAVII 420
QY 359 LSGDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEHVOL 418
Db 421 LSGDRPGLNVPRIEDIQNLQALELQKLHPRESSQLFAKLQKMTDLRQIVTEHVOL 480
QY 419 LOYIKKTETDMSLHPLLOEIRYKDLX 443
Db 481 LOYIKKTETDMSLHPLLOEIRYKDLX 505
```

```
RESULT 3
US-10-109-886-6
; Sequence 6, Application US/10109886
; Patent No. US20020119499A1
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND
; FILE REFERENCE: TANIUCHI-6
; CURRENT APPLICATION NUMBER: US/10/109,886
; PRIOR FILING DATE: 2002-04-01
; PRIOR APPLICATION NUMBER: 09/514,247
; PRIOR FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 6
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-109-886-6
```

```
Query Match 62.6%; Score 2196.5; DB 12; Length 506;
Best Local Similarity 86.8%; Pred. No. 2.7e-151;
Matches 439; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

QY 1 MGETLGDSPIDPESDSTDTLSANISOEMTMVTEMPFMTNFGISSVDLSVMEHSHSF 60
Db 1 MGETLGDSPIDPESDSTDTLSANISOEMTMVTELAFTNFGISSVDLSVMEHSHSF 60
QY 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVVADYKYDKLQEOYSAIKVEBPASPPYSEKT 120
Db 61 DIRPFTTVDFSSISTPHYEDIPFTRTDPVVADYKYDKLQEOYSAIKVEBPASPPYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKPHEEPSNLSMAIECVCGDKASGFHYGVHACBGCGFFRRIRIRLKIYDRCDLNC 180
```







```

; CURRENT APPLICATION NUMBER: US/09/765,111A
; CURRENT FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 27
; LENGTH: 478
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-765-111A-27

Query Match
Best Local Similarity 58.6%; Score 2054.5; DB 10; Length 478;
Matches 411; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

OY 29 MFWDTMPFPTNFGISSVDLSVMDHSHSPDIKPTTVDFSSISTPHYEDIPTRTDP 88
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MFWDEIAFWPTNFGISSVDLSVMDHSHSPDIKPTTVDFSSISTPHYEDIPTRTDP 60
OY 89 VVADYKIDLQEOYSAIKVEPASPYYSEKTLQYN----- 124
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 VVADYKIDLQEOYSAIKVEPASPYYSEKTLQYNPHEEPSNSLMAIECRVCGDKASG 120
OY 125 -----RNNCOYCRFOKCLAVGMSHNAI 146
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 FHVGHACGCGKGFRRTRILKLIYDRCDLNCRIHKSRNKCQYCRFOKCLAVGMSHNAI 180
OY 147 RFRMPQAEKEKLLAEISSDIDQNPESADLR-ALAKHLVDSYIKSFPLTKAKARAILTG 205
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 RFRRIQAEKEKLLAEISSDIDQNPESADLRQALAKHLVDSYIKSFPLTKAKARAILTG 240
OY 206 KTDKSPFYIYDNNSLMGEDIKFKHITPLOQSKVARIIFOGCQFRSVEAVOETTEY 265
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 KTDKSPFYIYDNNSLMGEDIKFKHITPLOQSKVARIIFOGCQFRSVEAVOETTEY 300
OY 266 AKSIPGFVNLNDQYTLKYGVHEIITMTLASIMNKDGVLSSEGQFMTREFLSLRKP 325
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 AKSIPGFVNLNDQYTLKYGVHEIITMTLASIMNKDGVLSSEGQFMTREFLSLRKP 360
OY 326 FGFMPKREFFAVKFNALELSDSLAIFIAVITLSDRGLNVKPIEDIONMLQALEL 385
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FGFMPKREFFAVKFNALELSDSLAIFIAVITLSDRGLNVKPIEDIONMLQALEL 420
OY 386 QLKLNPESSQLFAKLQKMTDLROIIVTEHVOLLQYIKKTEFDMSHPLLOETIKDLY 443
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 QLKLNPESSQLFAKLQKMTDLROIIVTEHVOLLQYIKKTEFDMSHPLLOETIKDLY 478

RESULT 9
US-10-142-373-2
; Sequence 2, Application US/10142373
; Patent No. US2002013765A1
; GENERAL INFORMATION:
; APPLICANT: EVANS, Ronald
; APPLICANT: FORMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED
; TITLE OF INVENTION: RECEPTOR-GAMMA, FOR THE USE THEREOF
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/10/142,373
; CURRENT FILING DATE: 2002-05-08
; PRIOR APPLICATION NUMBER: US/09/788,070
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 475
; TYPE: PRT
```

```

; ORGANISM: Mus Musculus
US-10-142-373-2

Query Match
Best Local Similarity 58.6%; Score 2054; DB 12; Length 475;
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

OY 31 MVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPTTVDFSSISTPHYEDIPTRTDPVY 90
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 1 MVDTEMPFPTNFGISSVDLSVMDHSHSPDIKPTTVDFSSISARHYEDIPTRTADPVY 60
OY 91 ADYKIDLQEOYSAIKVEPASPYYSEKTLQYN----- 124
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 61 ADYKIDLQEOYSAIKVEPASPYYSEKTLQYNRHEEPSNSLMAIECRVCGDKASGFH 120
OY 125 -----RNNCOYCRFOKCLAVGMSHNAIR 148
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 121 YGVHACGCGKGFRRTRILKLIYDRCDLNCRIHKSRNKCQYCRFOKCLAVGMSHNAIR 180
OY 149 GRMPQAEKEKLLAEISSDIDQNPESADLRALAKHLVDSYIKSFPLTKAKARAILTGKTT 208
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 181 GRMPQAEKEKLLAEISSDIDQNPESADLRALAKHLVDSYIKSFPLTKAKARAILTGKTT 240
OY 209 DKSPEVIYDNNSLMGEDIKFKHITPLOQSKVARIIFOGCQFRSVEAVOETTEYAKS 268
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 241 DKSPEVIYDNNSLMGEDIKFKHITPLOQSKVARIIFOGCQFRSVEAVOETTEYAKN 300
OY 269 IPEGFVNLNDQYTLKYGVHEIITMTLASIMNKDGVLSSEGQFMTREFLSLRKPFQD 328
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 301 IPEGFVNLNDQYTLKYGVHEIITMTLASIMNKDGVLSSEGQFMTREFLSLRKPFQD 360
OY 329 FMEKREFFAVKFNALELSDSLAIFIAVITLSDRGLNVKPIEDIONMLQALELQK 388
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 361 FMEKREFFAVKFNALELSDSLAIFIAVITLSDRGLNVKPIEDIONMLQALELQK 420
OY 389 LNHPESSQLFAKLQKMTDLROIIVTEHVOLLQYIKKTEFDMSHPLLOETIKDLY 443
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 421 LNHPESSQLFAKLQKMTDLROIIVTEHVOLLQYIKKTEFDMSHPLLOETIKDLY 475

RESULT 10
US-10-013-807-2
; Sequence 2, Application US/10013807
; Patent No. US20020112256A1
; GENERAL INFORMATION:
; APPLICANT: Guenther, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING
; TITLE OF INVENTION: PROLIFERATOR-ACTIVATED RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-683
; CURRENT APPLICATION NUMBER: US/10/013,807
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,916
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-013-807-2

Query Match
Best Local Similarity 30.7%; Score 1076.5; DB 12; Length 440;
Matches 203; Conservative 58; Mismatches 64; Indels 1; Gaps 1;

OY 119 KTDLYNRNKCQYCRFOKCLAVGMSHNAIRFGMPQAEKEKLLAEI-SSDIDQNPESADL 177
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 115 KTDLYNRNKCQYCRFOKCLAVGMSHNAIRFGMPQAEKEKLLAEI-SSDIDQNPESADL 174
OY 178 RALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPEVIYDNNSLMGEDIKFKHITPQ 237
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

Db 175 KAFSKHIYNAVLKFNMTKKKARSILTGKSSHNAPEVIHDIETLMOAKRGLWKQVLNGL 234  
QY 238 EOSEVAIRIFOGCGFSEVAEVOITEYAKSIPGFVNLNDVOYTLTKYGHETITMMLA 297  
Db 235 PPVNEISVHFYRQSTVEVRELTEFAKNIPNFSLSPLDVOYTLTKYGHETITMMLA 294  
QY 298 SLMNKDGVLISEGQGFMTREFLKSIRKPFQGFMPKFEFAVKFNALIEDSDLAIFIAVI 357  
Db 295 SIYKDKGLIVANGSGFYTHEFLRSRKPFSDIIEPKFEFANFKFALIEDSDLAIFIAI 354  
QY 358 ILSGDRPGLNVKPIEDIODNLQALQKLNHPSSQLEFAKLQKMTDIOIETVHQ 417  
Db 355 ILGDRPGLNVKPIEDIODNLQALQKLNHPSSQLEFAKLQKMTDIOIETVHQ 414  
QY 418 LLOVIRKTEEDMSLHPLLOETIKDLY 443  
Db 415 MMOWLKTESETLHPLLOETIKDLY 440

## RESULT 11

US-09-925-299-1029  
; Sequence 1029, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08  
; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 1029  
; LENGTH: 121  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (9)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; LOCATION: (108)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-1029

Query Match 14.9%; Score 524; DB 10; Length 121;  
Best Local Similarity 91.4%; Pred. No. 2,8e-31;  
Matches 106; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 462 DNMSLMGGEKIRFKHTITPLQEOSEKVAIRIFOGCGFSEVAEVOITEYAKSIPGFVNL 521  
Db 1 DNMSLMGGEKIRFKHTITPLQEOSEKVAIRIFOGCGFSEVAEVOITEYAKSIPGFVNL 60  
QY 522 LNDVOYTLTKYGHETITMMLASLMNKDGLVISEGQGFMTREFLKSIRKPFQGFMP 577  
Db 61 LNDVOYTLTKYGHETITMMLASLMNKDGLVISEGQGFMTREFLKSIRKPFQGFMP 116

## RESULT 12

US-09-925-299-882  
; Sequence 882, Application US/09925299  
; Patent No. US20020055627A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies  
; FILE REFERENCE: PA102  
; CURRENT APPLICATION NUMBER: US/09/925,299  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: PCT/US00/05883  
; PRIOR FILING DATE: 2000-03-08

; PRIOR APPLICATION NUMBER: 60/124,270  
; PRIOR FILING DATE: 1999-03-12  
; NUMBER OF SEQ ID NOS: 1556  
; SOFTWARE: Patentln Ver. 2.0  
; SEQ ID NO 882  
; LENGTH: 128  
; TYPE: PRF  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (96)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (112)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-882

Query Match 14.6%; Score 512.5; DB 10; Length 128;  
Best Local Similarity 85.2%; Pred. No. 2e-30;  
Matches 98; Conservative 2; Mismatches 4; Indels 11; Gaps 1;

QY 22 SANISQ-----EMTWDTMPFMPNFGISSVDSVMSHSHSFDIKPFTTVDF 70  
Db 1 TTNIQGHFLKRESAFNEMTMDTEMPFPTNFGISSVDSVMSHSHSFDIKPFTTVDF 60  
QY 71 SSISTPHYEDIPFTTRDPVADVADKYDKLOEYOSAIRVEBPAPPYSEKTOYLN 125  
Db 61 SSISTPHYEDIPFTTRDPVADVADKYDKLOEYOSAIRVEBPAPPYSEKTOYLN 115

## RESULT 13

US-10-108-605-213  
; Sequence 213, Application US/10108605  
; Patent No. US20020160934A1  
; GENERAL INFORMATION:  
; APPLICANT: Broadus, Julie  
; APPLICANT: Stam, Lynn  
; APPLICANT: Bachmann, Jane  
; APPLICANT: Kamdar, Kim  
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCOD  
; FILE REFERENCE: 31133B  
; CURRENT APPLICATION NUMBER: US/10/108,605  
; CURRENT FILING DATE: 2002-03-27  
; PRIOR APPLICATION NUMBER: US 09/761,142  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/176,418  
; PRIOR FILING DATE: 2000-01-14  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: Patentln Ver. 2.1  
; SEQ ID NO 213  
; LENGTH: 1394  
; TYPE: PRF  
; ORGANISM: Drosophila melanogaster  
US-10-108-605-213

Query Match 11.2%; Score 392; DB 9; Length 1394;  
Best Local Similarity 27.8%; Pred. No. 2,6e-20;  
Matches 121; Conservative 80; Mismatches 160; Indels 74; Gaps 16;

QY 48 VDSVMEHSHSFDIKPFTTVDFSSISTPHYEDIPFTTRDPV-----ADYKDLKQ-- 100  
Db 382 LDSQNLVLCQKDFDEKIDTALSNSSANT-----GRNTPAVTANEDADGFRRSIQOK 433  
QY 101 -EYOSAIRVEBPAPPYSEKTOYLNRRKCOYCFORCLAVGSMHAIREGRMPOAKEKL 159  
Db 434 IYRPTCKNQC-----ILRINRRCOYCRUKCIAVGSMDAVREGRVKREKARI 486  
QY 160 LAETSSIIDLNESADRLAKHLVD-----SYKSPPLTKAKARALLTGKTTD 209  
Db 487 LAA-----MOOSTONRGQORLATELDQPRLLAVALRAHLEDEFTKEKVSARORARD 542  
QY 210 KSPFVIYDMNSLMGGEKIRFKHTITPLQEOSEKVAIRIFOGCGFSEVAEVOITEYAKSI 269

```
Db 543 PS-----YSMPTLLACP-----LNPAPELQSE-----OEFSGREAHVIRGVIDFAGMI 585
OY 270 PGFVNLDLNOQVTLTKGYVEHIIYTMLASLNNK--DGVLLSEGGFTRFELKS-LKKPF 326
Db 586 PGFOLLQDDKRTKGLDPALEVRLLICMFDDSSINIIICLNCQ-VRRRAIIONGANARF 644
OY 327 GDFMEPFEEFAVFNALDSDLAIFIAVIIISGDRPGLNVPEDIODNLQALELQ 386
Db 645 --LVDSIFENFAERNNSNMLDAELGALCALVILTPDRPGLRNLELEKMSRLKGCIOYI 702
OY 387 LKLNHPSSQLFAKLQKMTDLROIVTEHVQLLQVTKTETMSLHPLEIYKDYAMA 446
Db 703 VAONRRPQPEFLAKLLETMPDLRTLSTLHTEKLVFTEHEKEL-----LROOM-----WS 752
OY 447 ILGKTTD---KSP 457
Db 753 MEDGNNSDGOONKSP 767

RESULT 14
US-10-108-605-211
: Sequence 211, Application US/10108605
: Patent No. US20020160934A1
: GENERAL INFORMATION:
: APPLICANT: Broadus, Julie
: APPLICANT: Stam, Lynn
: APPLICANT: Bachmann, Jane
: APPLICANT: Kamdar, Kim
: TITLE OF INVENTION: NUCLEIC ACID SEQUENCES FROM DROSOPHILA MELANOGASTER THAT ENCODE
: FILE REFERENCE: 31133B
: CURRENT APPLICATION NUMBER: US/10/108,605
: PRIOR FILING DATE: 2002-03-27
: PRIOR APPLICATION NUMBER: US 09/761,142
: PRIOR FILING DATE: 2001-01-16
: PRIOR APPLICATION NUMBER: US 60/176,418
: PRIOR FILING DATE: 2000-01-14
: NUMBER OF SEQ ID NOS: 361
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 211
: LENGTH: 1237
: TYPE: PRT
: ORGANISM: Drosophila melanogaster
US-10-108-605-211

Query Match 11.0%; Score 385.5; DB 9; Length 1237;
Best Local Similarity 29.3%; Pred. No. 6.4e-20;
Matches 103; Conservative 71; Mismatches 126; Indels 51; Gaps 12;

OY 124 NNRKCOYCFQKCLAVGMSHNAIRFGMPQAEKEKLLAEISSIDOLNPESADLRALAKH 183
Db 294 NNRKCOYCRKCKCIANGMSDAVRFQGVPRKREARILAA-----MOOSTORGOORALATE 349
OY 184 LND-----SYTKSPPLYAKARAILTGKTTDKSPFVIYDMSLMMGEDKIRFKHT 233
Db 350 LDDQPRLLAAVLAHLTECFETKEKVSAMRORADCP-----YSMPTLLACP-----L 398
OY 234 TPLQESKEVAIRIFQCGCFRSEVAOETEVAKSTPGFVNLDLNOQVTLTKGYVEHIIY 293
Db 399 NPAPPELQSE-----OEFSGREAHVIRGVIDFAGMI PGOLLQDDKFTLLKGLDPALE 452
OY 294 TMLASLNNK--DGVLLSEGGFTRFELKS-LKKPF GDFMEPFEEFAVFNALDSDLDL 350
Db 453 VRLICMFDDSSINIIICLNCQ-VRRRAIIONGANARF--LVDSIFENFAERNNSNMLDAEL 509
OY 351 AIFIAVIIISGDRPGLNVPEDIODNLQALELQ LKLNHPSSQLFAKLQKMTDLRO 410
Db 510 GLFCALVILTPDRPGLRNLELEKMSRLKGCIOYI VAONRRPQPEFLAKLLETMPDLRL 569
OY 411 IYVEHVQLLQVTKTETMSLHPLEIYKDYAMAILGKTTD---KSP 457
Db 570 LSTLHTEKLVFTEHEKEL-----LROOM-----WSMEDGNNSDGOONKSP 610
```

```
RESULT 15
US-09-864-761-35611
: Sequence 35611, Application US/09864761
: Patent No. US20020048763A1
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME- DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
: FILE REFERENCE: Aeomica-x-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: PRIOR FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263.6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00667
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00664
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00669
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00665
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00668
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00663
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00662
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00661
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: PCT/US01/00670
: PRIOR FILING DATE: 2001-01-30
: PRIOR APPLICATION NUMBER: US 60/234,687
: PRIOR FILING DATE: 2000-09-21
: PRIOR APPLICATION NUMBER: US 09/608,408
: PRIOR FILING DATE: 2000-06-30
: PRIOR APPLICATION NUMBER: US 09/774,203
: PRIOR FILING DATE: 2001-01-29
: NUMBER OF SEQ ID NOS: 49117
: SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
: SEQ ID NO 35611
: LENGTH: 81
: TYPE: PRT
: ORGANISM: Homo sapiens
: FEATURE:
: OTHER INFORMATION: MAP TO NM_007104.1
: OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
: OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 3.6
: OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.4
: OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
: OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.9
: OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 3
: OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
: OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.9
: OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 4
: OTHER INFORMATION: EST_HUMAN HIT: AU131142.1, EVALU6 5.00e-39
: OTHER INFORMATION: SWISSPROT HIT: Q03181, EVALU6 6.00e-40
US-09-864-761-35611
```



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: February 25, 2003, 03:58:31 ; Search time 22 Seconds  
(without alignments)  
920.135 Million cell updates/sec

Title: US-09-931-007A-1

Perfect score: 3508

Sequence: 1 MERTIGDSPIDPESDSFTDT.....KFTDMSLHPLLQETIKDLY 688

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/1aa/PTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	505	4	US-09-128-142-4
2	2196.5	62.6	506	4	US-09-514-247A-6
3	2089	59.5	477	4	US-09-128-142-2
4	2054	58.6	475	2	US-08-484-200-2
5	2054	58.6	475	3	US-08-465-375-2
6	2054	58.6	475	4	US-09-788-070-2
7	2044.5	58.3	476	4	US-08-134-557D-2
8	2043	58.2	475	4	US-08-764-870-10
9	2043	58.2	475	4	US-08-980-115-9
10	1101.5	31.4	441	4	US-08-764-870-9
11	1101.5	31.4	441	4	US-08-980-115-9
12	1081.5	30.8	468	1	US-08-459-287-3
13	1076.5	30.7	440	2	US-08-484-200-4
14	1076.5	30.7	468	1	US-08-459-287-2
15	1074.5	30.6	468	4	US-08-764-870-8
16	1074.5	30.6	468	4	US-08-980-115-8
17	1071	30.5	439	1	US-08-333-358-10
18	1071	30.5	439	1	US-08-463-694-10
19	1071	30.5	439	1	US-08-694-501-10
20	375.5	10.7	483	4	US-09-106-194-4
21	341.5	9.7	548	1	US-08-333-358-2
22	341.5	9.7	548	1	US-08-463-694-2
23	341.5	9.7	548	1	US-08-694-501-2
24	341.5	9.7	556	4	US-09-106-194-3
25	324.5	9.3	368	6	5223606-3
26	324.5	9.3	448	6	5223606-2
27	317	9.0	416	4	US-08-764-870-4

28	317	9.0	416	4	US-08-980-115-4	Sequence 4, Appli
29	316	9.0	462	6	5171671-2	Patent No. 5171671
30	315	9.0	462	2	US-08-592-383-2	Sequence 2, Appli
31	315	9.0	462	2	US-08-095-728B-4	Sequence 4, Appli
32	315	9.0	462	5	PCT-US92-02320A-4	Sequence 2, Appli
33	315	9.0	462	5	PCT-US92-02320A-4	Sequence 2, Appli
34	315	9.0	797	5	US-08-095-728B-2	Sequence 2, Appli
35	313	8.9	454	4	PCR-US92-02320A-2	Sequence 5, Appli
36	313	8.9	454	4	US-08-764-870-5	Sequence 5, Appli
37	313	8.9	454	4	US-08-980-115-5	Sequence 5, Appli
38	305	8.7	440	1	US-08-333-358-8	Patent No. 5260432
39	305	8.7	440	1	US-08-463-694-8	Sequence 8, Appli
40	305	8.7	440	1	US-08-694-501-8	Sequence 8, Appli
41	305	8.7	447	1	US-08-373-935-1	Sequence 1, Appli
42	296	8.4	557	4	US-08-653-648A-12	Sequence 12, Appli
43	293.5	8.4	403	2	US-08-592-383-4	Sequence 4, Appli
44	283.5	8.1	674	4	US-08-653-648A-14	Sequence 14, Appli
45	283	8.1	433	2	US-08-466-120-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1  
US-09-128-142-4  
Sequence 4, Application US/09128142  
Patent No. 6284559  
GENERAL INFORMATION:  
Applicant: Smith, Roy G.  
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND G  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: Jack L. Tribble  
STREET: 126 E. Lincoln Ave., P.O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: Power Macintosh 7500/100  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/128,142  
FILING DATE: 03-Aug-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/844,007  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Tribble, Jack L.  
REGISTRATION NUMBER: 32,633  
REFERENCE/DOCKET NUMBER: <Unknown>  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-5321  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 505 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 4:  
US-09-128-142-4  
Query Match 63.6%; Score 2231; DB 4; Length 505;  
Best Local Similarity 87.7%; Pred. No. 7.6e-186;  
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;  
0Y 1 MERTIGDSPIDPESDSFTDTLSANISQEMTWDTMPFMTNFGISSVDLSVMEHSHSF 60

```
|||||
Db 1 MGETLDSPLDPSDSFTDTLSANISQEMTMVDTMPFWMTNNGISSVDLSVMDHSHSF 60
QY 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVYADKYDKLQEOYSALKVEPASPYYSEKT 120
Db 61 DLRPFTTVDFSSISTPHYEDIPFTRTDPVYADKYDKLQEOYSALKVEPASPYYSEKT 120
QY 121 QLYN----- 124
Db 121 QLYNKRHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRITRLKLIYDRCDLNC 180
QY 125 -----RNCQYCRFOKCLAVGSHNAIRFGMPQAEKEKLAIESSDIDQLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOKCLAVGSHNAIRFGMPQAEKEKLAIESSDIDQLNPESADLR 240
QY 179 ALAKHLIDSYIKSFPLTKAKARAILTGKTTDKSPVYIDMNSLMGDEKIKFKHITPLOE 238
Db 241 ALAKHLIDSYIKSFPLTKAKARAILTGKTTDKSPVYIDMNSLMGDEKIKFKHITPLOE 300
QY 239 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLKYGVEIITYTMLAS 298
Db 301 OSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLKYGVEIITYTMLAS 360
QY 299 LMKKDGVLISEGQGFMTREFLKLKRPFGDPEMPKFEFAVKFNALDELDDSLAIFIAVII 358
Db 361 LMKKDGVLISEGQGFMTREFLKLKRPFGDPEMPKFEFAVKFNALDELDDSLAIFIAVII 420
QY 359 LSGDRGGLNVPRIEDIDQNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVQ 418
Db 421 LSGDRGGLNVPRIEDIDQNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVQ 480
QY 419 LQVIRKTEETDMSLHPLOEITYKDL 443
Db 481 LQVIRKTEETDMSLHPLOEITYKDL 505

RESULT 2
US-09-514-247A-6
; Sequence 6, Application US/09514247A
; Patent No. 6365361
; GENERAL INFORMATION:
; APPLICANT: TANABE SEIYAKU CO. LTD.
; APPLICANT: TANIGUCHI, Tomoyasu
; APPLICANT: MIZUKAMI, Junko
; TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO PPA
; FILE REFERENCE: TANIGUCHI-6
; CURRENT APPLICATION NUMBER: US/09/514,247A
; CURRENT FILING DATE: 2000-02-28
; PRIOR APPLICATION NUMBER: PCT/JP98/03734
; PRIOR FILING DATE: 1998-08-24
; PRIOR APPLICATION NUMBER: JP231084/1997
; PRIOR FILING DATE: 1997-08-27
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO: 6
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-514-247A-6

Query Match 62.6%; Score 2196.5; DB 4; Length 506;
Best Local Similarity 86.8%; Pred. No. 7.6e-183;
Matches 439; Conservative 2; Mismatches 2; Indels 63; Gaps 2;
```

```
|||||
Db 121 QLYNKRHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFFRRITRLKLIYDRCDLNC 180
QY 125 -----RNCQYCRFOKCLAVGSHNAIRFGMPQAEKEKLAIESSDIDQLNPESADLR 178
Db 181 RIHKSRNKCQYCRFOKCLAVGSHNAIRFGMPQAEKEKLAIESSDIDQLNPESADLR 240
QY 179 -ALAKHLIDSYIKSFPLTKAKARAILTGKTTDKSPVYIDMNSLMGDEKIKFKHITPLO 237
Db 241 QALAKHLIDSYIKSFPLTKAKARAILTGKTTDKSPVYIDMNSLMGDEKIKFKHITPLO 300
QY 238 EQSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLKYGVEIITYTMLA 297
Db 301 EQSKEVAIRIFOGCQFRSVEAVQETETAKSIPGFVNLDLNDQVTLKYGVEIITYTMLA 360
QY 298 SLNKKDGVLISEGQGFMTREFLKLKRPFGDPEMPKFEFAVKFNALDELDDSLAIFIAVI 357
Db 361 SLNKKDGVLISEGQGFMTREFLKLKRPFGDPEMPKFEFAVKFNALDELDDSLAIFIAVI 420
QY 358 ILSGDRGGLNVPRIEDIDQNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVQ 417
Db 421 ILSGDRGGLNVPRIEDIDQNLQALELQKLNHPRESSQLEFALKLOKMTDLROIVTEHVQ 480
QY 418 LQVIRKTEETDMSLHPLOEITYKDL 443
Db 481 LQVIRKTEETDMSLHPLOEITYKDL 506
```

```
RESULT 3
US-09-128-142-2
; Sequence 2, Application US/09128142
; Patent No. 6294559
; GENERAL INFORMATION:
; APPLICANT: Smith, Roy C.
; TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
; PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA1 AND C
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Jack L. Tribble
; STREET: 126 E. Lincoln Ave., P.O. Box 2000
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: US
; ZIP: 07065-0907
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 7500/100
; SOFTWARE: Microsoft Word 6.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/128,142
; FILING DATE: 03-Aug-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/844,007
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Tribble, Jack L.
; REGISTRATION NUMBER: 32,633
; REFERENCE/DOCKET NUMBER: <Unknown>
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-5321
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 477 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-128-142-2
```

Query Match 59.5%; Score 2089; DB 4; Length 477;  
Best Local Similarity 87.0%; Pred. No. 1.5e-173;  
Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

29 MTWDTMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDP 88  
1 MTWDTMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDP 60  
89 VVADYKDKLQEOYSAIKVEPASPYYSEKTOLYN-----RNKCOYCRFOKCLAVGSHNAI 124  
61 VVADYKDKLQEOYSAIKVEPASPYYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASG 120  
125 -----RNKCOYCRFOKCLAVGSHNAI 146  
121 FHYGVHAGECGKGFRRIRLKLITDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAI 180  
147 RFGMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGK 206  
181 RFGMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGK 240  
207 TTDKSPFVIYDMSLMMGEDEKIKFKHITPLOEQSKEVAIRIFOGCQFVSVEAVOEITEYA 266  
241 TTDKSPFVIYDMSLMMGEDEKIKFKHITPLOEQSKEVAIRIFOGCQFVSVEAVOEITEYA 300  
267 KSIIPGEVNDLNDQVTLKYGVEHIIYTMLASLMMKDGVLISEGGFMTREFLSLRKPF 326  
301 KSIIPGEVNDLNDQVTLKYGVEHIIYTMLASLMMKDGVLISEGGFMTREFLSLRKPF 360  
327 GDMPEKFEFAVKFNALDELDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 386  
361 GDMPEKFEFAVKFNALDELDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 420  
387 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLQVYIKETEDMSLHPLLOEITYKDL 443  
421 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLQVYIKETEDMSLHPLLOEITYKDL 477

RESULT 4  
US-08-484-200-2  
; Sequence 2, Application US/08484200  
; Patent No. 5861274  
; GENERAL INFORMATION:  
; APPLICANT: EVANS, RONALD M.  
; APPLICANT: FORMAN, BARRY M.  
; APPLICANT: KLIEMER, STEVEN A.  
; APPLICANT: ONG, ESTELITA S.  
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PRETTY, SCHROEDER, BRUGGEMANN & CLARK  
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000  
; CITY: LOS ANGELES  
; STATE: CALIFORNIA  
; COUNTRY: USA  
; ZIP: 90071  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,200  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: REITER, STEPHEN E.  
; REGISTRATION NUMBER: 31,192  
; REFERENCE/DOCKET NUMBER: P41 9971  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 619-546-1995  
; TELEFAX: 619-546-9392  
; INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
; LENGTH: 475 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-484-200-2

Query Match 58.6%; Score 2054; DB 2; Length 475;  
Best Local Similarity 85.5%; Pred. No. 1.7e-170;  
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

31 MVDTEMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDPV 90  
1 MVDTEMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDPV 60  
91 ADYKDKLQEOYSAIKVEPASPYYSEKTOLYN-----RNKCOYCRFOKCLAVGSHNAIRF 124  
61 ADYKDKLQEOYSAIKVEPASPYYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120  
125 -----RNKCOYCRFOKCLAVGSHNAIRF 148  
121 YGVHAGECGKGFRRIRLKLITDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRF 180  
149 GRMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGKT 208  
181 GRMPQAEKEXLLAEISSDIDOLNPESADLRALAKHLVDSYIKSPPLTKAKARAILTGKT 240  
209 DKSPFVIYDMSLMMGEDEKIKFKHITPLOEQSKEVAIRIFOGCQFVSVEAVOEITEYAS 268  
241 DKSPFVIYDMSLMMGEDEKIKFKHITPLOEQSKEVAIRIFOGCQFVSVEAVOEITEYAN 300  
269 IPGFVNDLNDQVTLKYGVEHIIYTMLASLMMKDGVLISEGGFMTREFLSLRKPF 328  
301 IPGFVNDLNDQVTLKYGVEHIIYTMLASLMMKDGVLISEGGFMTREFLSLRKPF 360  
329 FMEPKFEFAVKFNALDELDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 388  
361 FMEPKFEFAVKFNALDELDSDLAIFIAVIIISGDRGLLNKPIEDIODNLQALELQ 420  
389 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLQVYIKETEDMSLHPLLOEITYKDL 443  
421 LKLNHPSSQLFAKLQKMTDLRQIVTEHVQLQVYIKETEDMSLHPLLOEITYKDL 475

RESULT 5  
US-08-465-375-2  
; Sequence 2, Application US/08465375A  
; Patent No. 6022897  
; GENERAL INFORMATION:  
; APPLICANT: Evans, Ronald M.  
; APPLICANT: Forman, Barry M.  
; TITLE OF INVENTION: SELECTIVE MODULATORS OF PEROXISOME  
; TITLE OF INVENTION: PROLIFERATOR ACTIVATED RECEPTOR-GAMMA, AND METHODS FOR THE  
; FILE REFERENCE: SALK1470-1  
; CURRENT APPLICATION NUMBER: US/08/465,375A  
; CURRENT FILING DATE: 1995-06-05  
; EARLIER APPLICATION NUMBER: 08/428,559  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO: 2  
; LENGTH: 475  
; TYPE: prt  
; ORGANISM: Mus musculus  
US-08-465-375-2

Query Match 58.6%; Score 2054; DB 3; Length 475;  
Best Local Similarity 85.5%; Pred. No. 1.7e-170;  
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

31 MVDTEMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDPV 90  
1 MVDTEMPFPTNFGISSVDLSVMDHSHSFDIKPPTVDFSSISIPHYEDIPPTRTDPV 60

Db 1 MVDEMPFWPTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPETRADPMV 60  
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN----- 124  
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120  
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148  
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180  
QY 149 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDSTIKSPPLTKAKARAILTGKTT 208  
Db 181 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDSTIKSPPLTKAKARAILTGKTT 240  
QY 209 DKSPFVIYDMNSLMMGDKIKFKHITPLOESKEVAIRIFOGCQFRSEVAEOETTEYAKS 268  
Db 241 DKSPFVIYDMNSLMMGDKIKFKHITPLOESKEVAIRIFOGCQFRSEVAEOETTEYAKN 300  
QY 269 IPGFVNDLNDQVTLTKYGVHEIITMLASLMNKDGYLISEGQCFMREFLKSIRKPFCD 328  
Db 301 IPGFVNDLNDQVTLTKYGVHEIITMLASLMNKDGYLISEGQCFMREFLKSIRKPFCD 360  
QY 329 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALELQLK 388  
Db 361 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALELQLK 420  
QY 389 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYKDL 443  
Db 421 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYKDL 475

RESULT 6  
US-09-788-070-2  
; Sequence 2, Application US/09788070  
; Patent No. 6413994

; GENERAL INFORMATION:  
; APPLICANT: EVANS, Ronald  
; APPLICANT: FORMAN, Barry  
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED RECEPTOR-GAMMA,  
; TITLE OF INVENTION: AND METHODS FOR THE USE THEREOF  
; FILE REFERENCE: SALK1480-2  
; CURRENT APPLICATION NUMBER: US/09/788, 070  
; CURRENT FILING DATE: 2001-02-16  
; PRIOR APPLICATION NUMBER: US 09/955, 302  
; PRIOR FILING DATE: 1999-02-22  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 475  
; TYPE: PRT  
; ORGANISM: Mus Musculus  
US-09-788-070-2

Query Match 58.6%; Score 2054; DB 4; Length 475;  
Best Local Similarity 85.5%; Pred. No. 1.7e-170;  
Matches 406; Conservative 4; Mismatches 3; Indels 62; Gaps 1;

QY 31 MVDEMPFWPTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPETRADPMV 90  
Db 1 MVDEMPFWPTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPETRADPMV 60  
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN----- 124  
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120  
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148  
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180  
QY 149 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDSTIKSPPLTKAKARAILTGKTT 208  
Db 181 GRMPOAKEKELIAEISSDIDOLNPESADLRALAKHLVDSTIKSPPLTKAKARAILTGKTT 240

QY 209 DKSPFVIYDMNSLMMGDKIKFKHITPLOESKEVAIRIFOGCQFRSEVAEOETTEYAKS 268  
Db 241 DKSPFVIYDMNSLMMGDKIKFKHITPLOESKEVAIRIFOGCQFRSEVAEOETTEYAKN 300  
QY 269 IPGFVNDLNDQVTLTKYGVHEIITMLASLMNKDGYLISEGQCFMREFLKSIRKPFCD 328  
Db 301 IPGFVNDLNDQVTLTKYGVHEIITMLASLMNKDGYLISEGQCFMREFLKSIRKPFCD 360  
QY 329 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALELQLK 388  
Db 361 FMEKFEFAKFNALIEDDSDLAIFIAVITISGDRPGLLVKPIEDIQDNLQALELQLK 420  
QY 389 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYKDL 443  
Db 421 LNHESQSLFAKLQKMTDLROIVTEHVOLLQVITKKTETMSLHPLLOEYKDL 475

RESULT 7  
US-08-134-557D-2

; Sequence 2, Application US/08134557D  
; Patent No. 6200802

; GENERAL INFORMATION:  
; APPLICANT: Greene, Marianne E.  
; APPLICANT: Blumberg, Bruce  
; TITLE OF INVENTION: Human Peroxisome Proliferator Activated  
; TITLE OF INVENTION: Receptor Gamma: Compositions and Methods  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Rockey, Milnamow & Katz, Ltd.  
; STREET: 2 Prudential Plaza, Suite 4700 180 N. Stetson  
; CITY: Chicago  
; STATE: IL  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/134, 557D  
; FILING DATE:  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Katz, Martin L.  
; REGISTRATION NUMBER: 25,011  
; REFERENCE/DOCKET NUMBER: ARCH:098  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-616-5400  
; TELEFAX: 312-616-5460  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 476 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-134-557D-2

Query Match 58.3%; Score 2044.5; DB 4; Length 476;  
Best Local Similarity 85.9%; Pred. No. 1.1e-169;  
Matches 409; Conservative 2; Mismatches 2; Indels 63; Gaps 2;

QY 31 MVDEMPFWPTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPETRADPMV 90  
Db 1 MVDEMPFWPTNFGISSVDLSVMEHSHSPDIKPFVTVDSSISAPHYEDIPETRADPMV 60  
QY 91 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYN----- 124  
Db 61 ADVKYDLKLOEYOSAIKVEBPASPPYSEKTOLYNRPHEEPSNSLMAIECRVCGDKASGFH 120  
QY 125 -----RNKCQYCRFOKCLAVGSMHNAIRF 148  
Db 121 YGVHACGCKGFFRRTIRLKLIVDRCDLNCRIHKSSRNKCQYCRFOKCLAVGSMHNAIRF 180



QY 149 GRMPOAEKELLAETSSDIDOLNPESADLR-ALAKHLVDSYIKSPFLTKAKARAILTGKT 207  
181 GRIAQAEKELLAETSSDIDOLNPESADLRQALAKHLVDSYIKSPFLTKAKARAILTGKT 240  
QY 208 TDKSPFYIDMNSLMNGEDKIKFKHITPILQEOSKEVAIRIFOGCQFRSVAVOETEVAK 267  
241 TDKSPFYIDMNSLMNGEDKIKFKHITPILQEOSKEVAIRIFOGCQFRSVAVOETEVAK 300  
QY 268 STPGFVNLNDQVTLTKYGVHEIITMNLASLNKKGVLISSEGQGMTREFLSLRKPPG 327  
301 STPGFVNLNDQVTLTKYGVHEIITMNLASLNKKGVLISSEGQGMTREFLSLRKPPG 360  
QY 328 DMEPEKFEFAVFNALDELSDLAIFIAVILISGDRPGLLNKVPIDIDONLLQALELOJ 387  
361 DMEPEKFEFAVFNALDELSDLAIFIAVILISGDRPGLLNKVPIDIDONLLQALELOJ 420  
QY 388 LNHPESSQLFAKLQKMTDLROIVTEHVQLQVTKTETDMSLHPLQEIYKDL 443  
421 LNHPESSQLFAKLQKMTDLROIVTEHVQLQVTKTETDMSLHPLQEIYKDL 476  
Db  
RESULT 8  
US-08-764-870-10  
Sequence 10, Application US/08764870  
Patent No. 6236946  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Thomas S  
APPLICANT: Baxter, John D  
APPLICANT: Fletterick, Robert J  
APPLICANT: Wagner, Richard L  
APPLICANT: Kushner, Peter J  
APPLICANT: Aprilletti, James W  
APPLICANT: West, Brian  
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand  
TITLE OF INVENTION: Binding Domains  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward  
STREET: Five Palo Alto Square, 3000 El Camino Real  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,870  
FILING DATE: 13-DEC-1996  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,540  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,543  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,606  
FILING DATE: 14-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakamura, Jackie N  
REGISTRATION NUMBER: 35,966  
REFERENCE/DOCKET NUMBER: UCAL-246/01US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)843-5000  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 475 amino acids  
type: amino acid  
STRANDEDNESS:

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-764-870-10  
Query Match 58.2%; Score 2043; DB 4; Length 475;  
Best Local Similarity 84.8%; Pred. No. 1.5e-169;  
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;  
QY 31 MVDTEMPFWPTNFGTSSVLSVMSHDSHSDIKPFTVDSISPHYDIPFTTRDPV 90  
Db 1 MVDTEMPFWPTNFGTSSVLSVMSHDSHSDIKPFTVDSISPHYDIPFTTRDPV 60  
QY 91 ADYKYLKLOEYOSAIAKYVPAAPYSEKTOLYN----- 124  
Db 61 ADYKYLKLOEYOSAIAKYVPAAPYSEKTOLYN----- 120  
QY 125 -----RNKOYCRFOKCLAVGSHNAIRF 148  
Db 121 YGVHACEGCKGFERRIRLKLTYDRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRF 180  
QY 149 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHLVDSYIKSPFLTKAKARAILTGKT 208  
Db 181 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHLVDSYIKSPFLTKAKARAILTGKT 240  
QY 209 DKSPEFYIDMNSLMNGEDKIKFKHITPILQEOSKEVAIRIFOGCQFRSVAVOETEVAKS 268  
Db 241 DKSPEFYIDMNSLMNGEDKIKFKHITPILQEOSKEVAIRIFOGCQFRSVAVOETEVAKS 300  
QY 269 IGFVNLNDQVTLTKYGVHEIITMNLASLNKKGVLISSEGQGMTREFLSLRKPPG 328  
Db 301 IGFVNLNDQVTLTKYGVHEIITMNLASLNKKGVLISSEGQGMTREFLSLRKPPG 360  
QY 329 FMEPEKFEFAVFNALDELSDLAIFIAVILISGDRPGLLNKVPIDIDONLLQALELOJ 388  
Db 361 FMEPEKFEFAVFNALDELSDLAIFIAVILISGDRPGLLNKVPIDIDONLLQALELOJ 420  
QY 389 LNHPESSQLFAKLQKMTDLROIVTEHVQLQVTKTETDMSLHPLQEIYKDL 443  
Db 421 LNHPESSQLFAKLQKMTDLROIVTEHVQLQVTKTETDMSLHPLQEIYKDL 475  
RESULT 9  
US-08-980-115-10  
Sequence 10, Application US/08980115  
Patent No. 6266622  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Thomas S.  
APPLICANT: Baxter, John D.  
APPLICANT: Fletterick, Robert J.  
APPLICANT: Wagner, Richard L.  
APPLICANT: Kushner, Peter J.  
APPLICANT: Aprilletti, James W.  
APPLICANT: West, Brian L.  
APPLICANT: Shiau, Andrew K.  
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS  
FILE REFERENCE: UCAL-246/02US  
CURRENT APPLICATION NUMBER: US/08/980,115  
FILING DATE: 1997-11-26  
EARLIER APPLICATION NUMBER: 08/764,870  
EARLIER FILING DATE: 1996-12-13  
EARLIER APPLICATION NUMBER: 60/008,606  
EARLIER FILING DATE: 1995-12-14  
EARLIER APPLICATION NUMBER: 60/008,543  
EARLIER FILING DATE: 1995-12-13  
EARLIER APPLICATION NUMBER: 60/008,540  
EARLIER FILING DATE: 1995-12-13  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: Patentln Ver. 2.0  
SEQ ID NO 10  
LENGTH: 475  
type: PRT  
ORGANISM: Homo sapiens  
FEATURE:

NAME/KEY: DOMAIN  
LOCATION: (202) (475)  
OTHER INFORMATION: minimal ligand binding domain  
US-08-980-115-10

Query Match 58.2%; Score 2043; DB 4; Length 475;  
Best Local Similarity 84.8%; Pred. No. 1.5e-169;  
Matches 403; Conservative 6; Mismatches 4; Indels 62; Gaps 1;

```

QY 31 MVDTEMPPTNGISSVSLVSMEDSHSFDIRKPTTVDFSSISPTPHYDIDPTPTDPV 90
DB 1 MVDTEMPPTNGISSVSLVSMEDSHSFDIRKPTTVDFSSISADPHEDIPTRADPV 60
QY 91 ADVKYDLKQEOASAKVBPASPPYSEKTOLYN----- 124
DB 61 ADVKYDLKQEOASAKVBPASPPYSEKTOLYNRPHEPSNSLMAIECRVCGDKASGFH 120
QY 125 -----RNKCQYCRFOKCLAVGSHNAIRF 148
DB 121 YGVHACEGCKGFEFRTIRKLIVRCDLNCRIHKSRNKCQYCRFOKCLAVGSHNAIRF 180
QY 149 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHYDSYIKSFPLTKAKARAILTGKT 208
DB 181 GRMPOAEKELLAETSSDIDOLNPESADLRALAKHLYDSITISFPLTKAKARAILTGKT 240
QY 209 DKSPFVYDMNSLMGEDIKFKHITPLOSKEVAIRIFOGCOFSESVAVOETFEYAKS 268
DB 241 DKSPFVYDMNSLMGEDIKFKHITPLOSKEVAIRIFOGCOFSESVAVOETFEYAKN 300
QY 269 IGFVNLIDNDVTLTKYGVHEITITMLASLNMKGCVLISEGOGFMTREFLSLRKPGD 328
DB 301 IGFVNLIDNDVTLTKYGVHEITITMLASLNMKGCVLISEGOGFMTREFLSLRKPGD 360
QY 329 FNEPKFEFAVFNALDSDLAIFAVITILSGDRPGLNVKPIEDIONLLOALELOK 388
DB 361 FNEPKFEFAVFNALDSDLAIFAVITILSGDRPGLNVKPIEDIONLLOALELOK 420
QY 389 LNHRESSQFAKLKMTDLROIYVTEHVOLLQYIKKTEFDMSLHPLLOEYKDLV 443
DB 421 LNHRESSQFAKLKMTDLROIYVTEHVOLLQYIKKTEFDMSLHPLLOEYKDLV 475

```

RESULT 10  
US-08-764-870-9  
Sequence 9, Application US/08764870  
Patent No. 6236946  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Thomas S  
APPLICANT: Baxter, John D  
APPLICANT: Fletterick, Robert J  
APPLICANT: Wagner, Richard L  
APPLICANT: Kushner, Peter J  
APPLICANT: Aprietti, James W  
APPLICANT: West, Brian  
TITLE OF INVENTION: Nuclear Receptor Ligands and Ligand  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooley Godward  
STREET: Five Palo Alto Square, 3000 El Camino Real  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94306  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,870  
FILING DATE: 13-DEC-1996  
CLASSIFICATION: 530

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,540  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,543  
FILING DATE: 13-DEC-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/008,606  
FILING DATE: 14-DEC-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Nakamura, Jackie N  
REGISTRATION NUMBER: 35,966  
REFERENCE/DOCKET NUMBER: UCAL-246/0105  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650)843-5000  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 441 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-764-870-9

Query Match 31.4%; Score 1101.5; DB 4; Length 441;  
Best Local Similarity 63.8%; Pred. No. 9.8e-88;  
Matches 208; Conservative 56; Mismatches 61; Indels 1; Gaps 1;

```

QY 119 KQLYRNKCQYCRFOKCLAVGSHNAIRFGRMPOAEKELLAETSSDIDOLNPESADL 177
DB 116 KQKRNKCQYCRFOKCLAVGSHNAIRFGRMPEAKRKLVAGLVANGSOYNQVADL 175
QY 178 RALAKHLYDSYIKSFPLTKAKARAILTGKTDPSPFVYDMNSLMGEDIKFKHITPQ 237
DB 176 KAFSKITVAYLKNFMTKKARSITGKASHTAPVINDITLMOAEGVLWKOLVNL 235
QY 238 EOSKEVAIRIFOGCOFSESVAVOETFEYAKSIPGFVNLIDNDVTLTKYGVHEITITMLA 297
DB 236 PPKKEISVHVFRCQTTVEYRELTEFAKSIPIFSFLNDQVTLTKYGVHEITITMLA 295
QY 298 SLNMKGCVLISEGOGFMTREFLSLRKPGDFMEKFEFAVFNALDSDLAIFIAI 357
DB 296 SIYMKDGLVANGSGFVTREFLSLRKPPSDIIEPKFEFAVFNALDSDLAIFIAI 355
QY 358 IISGDRPGLNVKPIEDIONLLOALELOKLNHRESSQFAKLKMTDLROIYVTEHV 417
DB 356 IISGDRPGLNVKPIEDIONLLOALELOKLNHRESSQFAKLKMTDLROIYVTEHV 415
QY 418 LLOVYIKKTEFDMSLHPLLOEYKDLV 443
DB 416 MMQRIKTEFETSLHPLLOEYKDMV 441

```

RESULT 11  
US-08-980-115-9  
Sequence 9, Application US/08980115  
Patent No. 6266622  
GENERAL INFORMATION:  
APPLICANT: Scanlan, Thomas S  
APPLICANT: Baxter, John D  
APPLICANT: Fletterick, Robert J  
APPLICANT: Wagner, Richard L  
APPLICANT: Kushner, Peter J  
APPLICANT: Aprietti, James W  
APPLICANT: West, Brian L  
APPLICANT: Shiao, Andrew K  
TITLE OF INVENTION: NUCLEAR RECEPTOR LIGANDS AND LIGAND BINDING DOMAINS  
FILE REFERENCE: UCAL-246/0205  
CURRENT APPLICATION NUMBER: US/08/980,115  
CURRENT FILING DATE: 1997-11-26  
EARLIER APPLICATION NUMBER: 08/764,870  
EARLIER FILING DATE: 1996-12-13  
EARLIER APPLICATION NUMBER: 60/008,606









GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

## OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 25, 2003, 04:00:36 ; Search time 69 seconds  
(without alignments)  
3057.878 Million cell updates/sec

Title: US-09-931-007A-1

Perfect score: 3508  
Sequence: 1 MGFTLGDSPIDPESDSFTDT.....KTETDMSLHPLQETRYDLY 688

## Scoring table:

BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlh  
-O=/cgn2\_1/USPTO/US09931007/runat\_14022003.102554.4697/app\_query.fasta\_1.839  
-DB=Issued\_Patents\_NA -QEXT=fastac -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MULEN=0 -MAXLEN=200000000  
-USER=US09931007.ecgn.1.1.32.runat.14022003.102554.4697 -NCP=6 -ICPU=3  
-NO\_XIPYX -NO\_MMAP -LARGQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN.TIMEOUT=30 -THRAIDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCITUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2231	63.6	1518	4	US-09-128-142-3
2	2231	63.6	1608	3	Sequence 3, Appli
3	2196.5	62.6	1679	4	US-09-484-345-3
4	2089	59.5	1434	4	US-09-514-247A-5
5	2065.5	58.9	1844	4	US-09-128-142-1
6	2065	58.9	1796	4	US-08-134-357D-1
7	2065	58.9	1796	4	US-09-255-392-1
8	2065	58.9	2005	2	US-08-484-200-1
9	2065	58.9	2005	2	US-08-477-493-1
10	2065	58.9	2005	3	US-08-465-375-1
11	1076.5	30.7	2012	4	US-09-788-070-1
12	1071	30.5	2009	2	US-08-484-200-3
					Sequence 9, Appli

13	1071	30.5	2009	1	US-08-463-694-9	Sequence 9, Appli
14	1071	30.5	2009	1	US-08-694-501-9	Sequence 9, Appli
15	1057.5	30.1	1407	1	US-08-459-287-1	Sequence 1, Appli
16	521	14.9	373	3	US-08-917-653-3	Sequence 3, Appli
17	341.5	9.7	1952	1	US-08-333-358-1	Sequence 3, Appli
18	341.5	9.7	1952	1	US-08-463-694-1	Sequence 1, Appli
19	341.5	9.7	1952	1	US-08-694-501-1	Sequence 1, Appli
20	333	9.5	277	3	US-08-917-653-4	Sequence 4, Appli
21	326.5	9.3	2989	6	523606-1	Patent No. 523606
22	321.5	9.2	3511	2	US-08-892-747-13	Sequence 13, Appli
23	315	9.0	2928	2	US-08-095-728B-3	Sequence 3, Appli
24	315	9.0	2928	5	PCT-US92-02320A-3	Sequence 3, Appli
25	315	9.0	2940	2	US-08-592-383-1	Sequence 1, Appli
26	315	9.0	2940	6	5171671-1	Patent No. 5171671
27	315	9.0	3036	1	US-08-306-691B-52	Sequence 52, Appli
28	315	9.0	3036	2	US-08-095-728B-1	Sequence 1, Appli
29	315	9.0	3036	5	PCT-US92-02320A-1	Sequence 1, Appli
30	313	8.9	1576	6	5260432-1	Patent No. 5260432
31	305	8.7	1659	1	US-08-333-358-7	Sequence 7, Appli
32	305	8.7	1659	1	US-08-463-694-7	Sequence 7, Appli
33	305	8.7	1659	1	US-08-694-501-7	Sequence 7, Appli
34	293.5	8.4	2658	2	US-08-592-383-3	Sequence 3, Appli
35	289.5	8.3	2043	1	US-07-737-736B-6	Sequence 6, Appli
36	283	8.1	1649	2	US-08-466-120-1	Sequence 1, Appli
37	283	8.1	1649	5	PCT-US94-07266-1	Sequence 1, Appli
38	281.5	8.0	1677	2	US-08-372-652-14	Sequence 14, Appli
39	281.5	8.0	1677	2	PCT-US95-16311-14	Sequence 14, Appli
40	278.5	7.9	1959	1	US-08-342-411A-3	Sequence 3, Appli
41	278.5	7.9	1959	5	PCT-US94-12883-4	Sequence 4, Appli
42	278	7.9	2231	1	US-08-496-631-1	Sequence 1, Appli
43	276.5	7.9	1662	1	US-08-336-408B-5	Sequence 5, Appli
44	276.5	7.9	1662	5	PCT-US91-00399-5	Sequence 5, Appli
45	276.5	7.9	1860	2	US-08-372-652-7	Sequence 7, Appli

## ALIGNMENTS

RESULT 1  
US-09-128-142-3  
Sequence 3, Application US/09128142  
Patent No. 6294559  
GENERAL INFORMATION:  
APPLICANT: Smith, Roy G.  
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH  
PEROXISOME PROLIFERATOR ACTIVATED RECEPTORS GAMMA AND G  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESS: 126 E. Lincoln Ave., P.O. Box 2000  
CITY: Rahway  
STATE: New Jersey  
COUNTRY: US  
ZIP: 07065-0907  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: Power Macintosh 7500/100  
SOFTWARE: Microsoft Word 6.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/128,142  
FILING DATE: 03-Aug-1998  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/844,007  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Tribble, Jack L.  
REGISTRATION NUMBER: 32,633  
REFERENCE/DOCKET NUMBER: <Unknown>  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 594-5321  
TELEFAX: (908) 594-4720  
INFORMATION FOR SEQ ID NO: 3:

;	SEQUENCE CHARACTERISTICS:	
;	LENGTH: 1518 base pairs	
;	TYPE: nucleic acid	
;	STRANDEDNESS: single	
;	TOPOLOGY: linear	
;	MOLECULE TYPE: DNA (genomic)	
;	SEQUENCE DESCRIPTION: SEQ ID NO: 3:	
;	US-09-128-142-3	
	Alignment Scores:	
	Pred. NO.:	6,98e-252
	Score:	2231.00
	Percent Similarity:	87.72%
	Best Local Similarity:	87.72%
	Query Match:	63.60%
	DB:	4
		1
	US-09-931-007A-1 (1-688) x US-09-128-142-3 (1-1518)	
0Y	1 MetGlyGurThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr	20
Db	1 ATGGGTGAACCTCTGGAGATTCTCTTATGACCAGAAAGCATTCCTTCACAGATACA	60
0Y	21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheIppro	40
Db	61 CTGTCTGCAACATATATCAAGAAATGACCACTGGTTACACAGAGATGCCATTCCTGGCCC	120
0Y	41 ThrAspPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe	60
Db	121 ACCAACTTTGGGATACAGCTCCGCTGATCTCTCCGTAATGGAAGACCACTCCACTCTT	180
0Y	61 AspIleIysProPheThrThrValAspPheSerIleSerThrProHisIlyrGluAsp	80
Db	181 GATATCAAGCCCTTCACACTGTTGACTTCCGACATTTCTACTCCACATTCACGAAGAC	240
0Y	81 IlePhePheThrArgThrAspProValValAlaAspIlyrIlyrAspLeuIlyrGluIn	100
Db	241 ATTCATTTCACAGAACGATTCAGTGGTTCACATTCACAGATTCACCGAAACTTCAC	300
0Y	101 GluIyrGlnSerAlaIleIyIysValGluProIleAspProProIyIyrSerGluIysThr	120
Db	301 GAGTACCAAAAGTCACATGAAGTGAGAGCTTCACATCTCATTTATTCGTGAAAGACT	360
0Y	121 GlnLeuIyrAsn-----	124
Db	361 CAGCTCTACATTAAGCCCTCATGAAGAGCTTCACAACTCCCTCATGCAATTGAATGCGT	420
0Y	124 -----	124
Db	421 GTCTGTGAGATTAAGCTTCGTGATTTCCACTATGAGAGCTTACGCTTGTGAAGATGCAG	480
0Y	124 -----	124
Db	481 GGTTCCTCCGAGAACATAGATTGAAGCTTATCTATATACAGAGTGAATCTTAACTGT	540
0Y	125 -----	124
Db	541 CGGATCCACAAAAAAGAGAAATAAATATGTCAGACTGTGGTTTTCAGAAATGCCTTTGA	600
0Y	139 ValGlyMetSerHisAsnAlaIleArgPheGlyIyrArgMetProGlnAlaGluIySGluIys	158
Db	601 GTGGGAGATGCTCATTAAGCTCAGTTCAGTTTGGCGGATGCCACAGGCGGAGAGGAAG	660
0Y	159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuIy	178
Db	661 CTGTGGGGGAGATCTCCAGGATATGACACCGCTGATCCAGAGTCCGCTGACCTCCGG	720
0Y	179 AlaLeuAlaIySHisLeuIyrAspSerTyrIleIySserPheProLeuThrIlyrAlaIys	198
Db	721 GCCCGGCAAAACATTTGTATGACTCATACATAAAGTCTTCCCGGTGACCAAGCAAG	780
0Y	199 AlaArgAlaIleLeuThrGlyIysThrThrAspIySserProPheValIleIyrAspMet	218

Alignment Scores:

```
Db 781 GGGAGGGGAGTCTTACACAGAAAGCAACAGACAAATACCATTCGTTATGTATGACATG 840
QY AsnSerLeuMetMetCylValuAspLysIleLysPheLysSHSIIeThrProLeuGlnIu 238
Db 841 AATTCCTTAATGATGGGAGAGATAAATCAAGTTCAAAACACATCACCCCTCGAGAG 900
QY GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
Db 901 CAGACCAAGAGAGGTGCCATCCGCACTTTTCAGGGCTCCAGGTTTCGCTCCGAGAGCT 960
QY ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
Db 961 GTGCGAGGAGATACAGAGATATGCCAAAGCAATCTCGGTTTTGTTAAATCTTGACTTGAC 1020
QY AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
Db 1021 GACCAAGTAACCTCTCTCAATATGTGACTCCACAGGATCATTTTACACATCTCGCCCTCC 1080
QY LeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPhe 318
Db 1081 TTGATGAATAAAGATGCGGGTTCTCATATCCAGGGGCCAAGGCTTCATGACAAAGGAGCTTT 1140
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
Db 1141 CTAAAGAGCGCTCGAAAGCCCTTTTGTGATCTTATGAGGCCCAAGTTGAAGTTGAGTTGCTGTG 1200
QY 339 LysPheAsnAlaLeuGluLeuAspSerSerAspLeuAlaIlePheIleValIleIle 358
Db 1201 AAGTTCATATGCACTGGATTTAGATGACAGCGACTTGGCAATATTTATTTGCTGCTATTATT 1260
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
Db 1261 CTCACGTGGAGACCGCCGAGGTTTGTCTGAATGTGAAGCCCATTTGAAGACATTCAGAGACAC 1320
QY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLysAsnHisProGluSerSerGluLeuPhe 398
Db 1321 CTGGTACAGAGCCCTGGAGGCTCAGCTGAAGCTGAACCAACCCCTGAGTCTCTCAGAGCTGTTT 1380
QY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
Db 1381 GCCAAGGTGCTCCAAATAATGACAGACCTCAGACAGATTTGACGGAACACAGTGCAGCTA 1440
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
Db 1441 CTGCGAGTGATCAAGAAAGACGAGACAGACATGATGATCTTCAACCCGCTCTCAGAGATTC 1500
QY 439 TyrLysAspLeuTyr 443
Db 1501 TTACAAGGACTGTGTAC 1515

RESULT 2
US-09-484-345-3
; Sequence 3, Application US/09484345
; Patent No. 6159734
; GENERAL INFORMATION:
; APPLICANT: Robert McKay
; APPLICANT: Alexander H. Borchers
; APPLICANT: Brenda F. Baker
; TITLE OF INVENTION: ANTISENSE MODULATION OF PEROXISOME PROLIFERATOR-ACTIVATED REE
; FILE REFERENCE: RFS-0104
; CURRENT APPLICATION NUMBER: US/09/484,345
; CURRENT FILING DATE: 2000-01-18
; NUMBER OF SEQ ID NOS: 90
; SEQ ID NO 3
; LENGTH: 1608
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)...(1608)
; US-09-484-345-3
```



Pred. No.: 7.65e-252 Length: 1608  
Score: 2231.00 Matches: 443  
Percent Similarity: 87.72% Conservative: 0  
Best Local Similarity: 87.72% Mismatches: 0  
Query Match: 63.60% Indels: 62  
DB: 3 Gaps: 1

US-09-931-007a-1 (1-688) x US-09-484-345-3 (1-1608)

QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
DB 91 AUGGGTGAACCTCGGGGATTCCTCTATTGACCCAGAAAGGATTCCTTCACTATACA 150  
QY 21 LeuSerAlaAsnIleSerGlnIleuMetThrMetValAspThrGluMetProPheThrPro 40  
DB 151 CTGTCTGCAAAATATTCACAGAAATGACCATGTTGACACAGAAATGCCATTCTGGCCC 210  
QY 41 ThrAsnPhgIlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
DB 211 ACCAATTGGGATCAGCTCCGATGATCTCTCCATATGGAAGACACCTCCACTCTT 270  
QY 61 AspIleIysProPheThrThrValAspPheSerSerIleSerThrProHisThrGluAsp 80  
DB 271 GATATCAAGCCCTTCACTACTGTTGACTTCCACATTCTACCTCCACATTACCAAGAAC 330  
QY 81 IleProPheThrArgThrAspProValAlaAspTyrIysTyrAspLeuIysLeuGln 100  
DB 331 ATTCATTACACAGAACAGATCCAGTGTGACAGATTCAAGATTAGACCTGAACCTTCAA 390  
QY 101 GluTyrGlnSerAlaIleIysValGluProAlaSerProProTyrTyrSerGluIysThr 120  
DB 391 GAGTACCAAGAGTCAATCAATGAAGTGAAGCTGATCTCCACCTTATTATTCTGAGAAGCT 450  
QY 121 GlnLeuTyrAsn----- 124  
DB 451 CAGCTCTCAATTAAGCCCTCATGAGAGCCCTTCCAACTCCCTCATGCAATTGAATGCTGCT 510  
QY 124 ----- 124  
DB 511 GTCGTGGAGATAAGCTTCGTGATTTCACTATGAGAGTTCACTTGTGAAGAGATGCAG 570  
QY 124 ----- 124  
DB 571 GGTTCCTCCGAGAACATCATGATTGAAGCTTATCTATGACAGATGATCTTAAGTCT 630  
QY 125 ----- ArgAsnIysCysGlnTyrCysArgPheGlnIysCysLeuAla 138  
DB 631 CGGATCCACAAAAAAGTAGAATTAATGTCAGTCTCGGTTTCAAAAAAGCCTTGCA 690  
QY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluIysGluIys 158  
DB 691 GTGGGGATGTCCTCATTAATGCCATCAGGTTTGGGGGATGCCACAGGCCGAGAGAGAG 750  
QY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
DB 751 CTGTGGCGGAGATCTCCAGTATGATGACACCAAGTGAATCCAGAGATCCGCTGACCTCGG 810  
QY 179 AlaLeuAlaIysHisLeuTyrAspSerTyrIleIysSerPheProLeuThrIysAlaIys 198  
DB 811 GCCCTGGCAAAACATTTGATGACTCATACATAAAGTCTTCCCTGCCGCAAAACAAAG 870  
QY 199 AlaArgAlaIleLeuThrGlyIysThrThrAspIysSerProPheValIleTyrAspMet 218  
DB 871 GCGAGGGGATCTTACAGAGAGAACACACAGACAAATCACCTTGTATCTATGACATG 930  
QY 219 AsnSerLeuMetMetGlyGluAspIysIleIysPheIysHisIleThrProLeuGlnIle 238  
DB 931 AATTCCTTAATGATGAGAGATTAATCAAGTCAACAAACATACACCCCTCGAGAGAG 990  
QY 239 GlnSerIysGluValAlaIleArgIlePheGlnIysCysGlnPheArgSerValGluAla 258  
DB 991 CAGAGCAAAAGAGTGGCCATCGCATCTTTTCAAGGGCTGCCAGTTTCCGCTCGTGAGGCT 1050

QY 259 ValGlnGluIleThrGluTyrAlaIysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
DB 1051 GTGCAGAGATCACAGAGATATGCCAAAAAGCATTCCTGCTTTGTAAATCTTGACTTGAC 1110  
QY 279 AspGlnValThrLeuLeuIysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298  
DB 1111 GACCAAGTAACTCTCCCAATATGAGATGCCAGAGATCATTTACACATGCTGCCCTCC 1170  
QY 299 LeuMetAsnIysAspGlyValLeuIleSerGluIysGlnIysPheMetThrArgGluPhe 318  
DB 1171 TTGATGAATTAAGATGGGGTCTCATATCCGAGGGCCAAAGCCTTATGACAAGGAGTTT 1230  
QY 319 LeuIysSerLeuArgIysProPheGlyAspPheMetGluProIysPheGluPheAlaVal 1338  
DB 1231 CTAAGAGACCTCCGAAAGCCTTTTGTGATCTTTATGAGGCCAAGTTTGAGTTCTGTG 1290  
QY 339 LysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
DB 1291 AAGTTCATGCACCTGGAATGATGATGACAGCGACTTGGCAATTTATGCTGTCAATT 1350  
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValIysProIleGluAspIleGlnAspAsn 378  
DB 1351 CTCAGTGGAGACCGCCAGGTTTCTGAATGGAAGCCCATTTGAAGACATTCAAAGAAC 1410  
QY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuIysLeuAsnHisProGluSerSerGlnLeuPhe 398  
DB 1411 CTGCTACAGCCCTGAGCTCCAGCTGGAAGCTGAACCCCTGAGTCTTCACAGCTGTTT 1470  
QY 399 AlaIysLeuLeuGlnIysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418  
DB 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTCAGGAAACGTCGAGCTA 1530  
QY 419 LeuGlnValIleIysIysThrGluThrAspMetSerLeuHisProLeuLeuGlnIle 438  
DB 1531 CTGCAGGTGATCAAGAAACGAGACAGACATGACTTTCACCCGCTCCTCGAGAGATC 1590  
QY 439 TyrIysAspLeuTyr 443  
DB 1591 TACAAAGACTTGTAC 1605

RESULT 3  
US-09-514-247a-5  
: Sequence 5, Application US/09514247a  
: Patent No. 6365361  
: GENERAL INFORMATION:  
: APPLICANT: TANABE SEIYAKU CO. LTD.  
: APPLICANT: TANIGUCHI, Tomoyasu  
: APPLICANT: MIZUKAMI, Junko  
: TITLE OF INVENTION: METHOD FOR IDENTIFYING OR SCREENING AGONIST AND ANTAGONIST TO  
: FILE REFERENCE: TANIGUCHI=6  
: CURRENT APPLICATION NUMBER: US/09/514,247a  
: CURRENT FILING DATE: 2000-02-28  
: PRIOR APPLICATION NUMBER: PCT/JP98/03734  
: PRIOR FILING DATE: 1998-08-24  
: PRIOR APPLICATION NUMBER: JP231084/1997  
: PRIOR FILING DATE: 1997-08-27  
: NUMBER OF SEQ ID NOS: 10  
: SOFTWARE: PatentIn version 3.0  
: SEQ ID NO 5  
: LENGTH: 1679  
: TYPE: DNA  
: ORGANISM: Homo sapiens  
: FEATURE:  
: NAME/KEY: CDS  
: LOCATION: (159)..(1679)  
US-09-514-247a-5

Alignment Scores:  
Pred. No.: 9.12e-248 Length: 1679  
Score: 2196.50 Matches: 439  
Percent Similarity: 87.15% Conservative: 2  
Best Local Similarity: 86.76% Mismatches: 2  
Query Match: 62.61% Indels: 63

```

DB:          4          Gaps:          2
US-09-931-007A-1 (1-688) x US-09-514-247A-5 (1-1679)
OY      1 MetGlyGlnThrLeuGlnLysSerProIleAspProGlnSerSerPheThrAspThr 20
        |||
DB      159 ATGGGGAAACCTGGGAAATTCCTATTGACCCAGAAAGCCCTTCACATGTATACA 218
OY      21 LeuSerAlaAsnIleSerGlnLumethrMetValAspThrGlnLumethrPro 40
        |||
DB      219 CTGTGCAAAACATATACAAAGAAATGACCATGTTGACACAGAGATGCCATTTCTGCCC 278
OY      41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGlnAspHisSerHisSerPhe 60
        |||
DB      279 ACCAACTTGGGATGAGTCGCGGATCTCTCGTAAATGGAAGACCACTCCACCTCTTT 338
OY      61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisThrLysAsp 80
        |||
DB      339 GATATCAAGCCCTTCCTACTCTGTGACTCTCCAGCATTTCACTCCACATTCAGCAAGAC 398
OY      81 IleProPheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysLeuGln 100
        |||
DB      399 ATTCCTATTACAAAGAACATCCAGTGGTTGCAGATTACAGATGATGACCTGAACCTTCAA 458
OY      101 GlnTyrGlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGlnLysThr 120
        |||
DB      459 GAGTACCAAGGTCATCAATCAAGTGGAGCTGCATCTCACCTTATATATCTGAGAGACT 518
OY      121 GlnLeuTyrAsn----- 124
DB      519 CAGCTCTACAAATTAAGCCTCATGAAGCCTTCCAACTCCCTCATGGCAAATTGAATGTCGT 578
OY      124 ----- 124
DB      579 GTCGTGAGATAAAGCTTCTGGATTTCACTATGGAGTTCATGCTTGTAAGATCCAG 638
OY      124 ----- 124
DB      639 GGTTTCTTCGGAAGAAATCAATGAACTTATCTATGACAGATGTGATCTTAAGTCT 698
OY      125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
        |||
DB      699 CGGATCCACAAAAAGTGAATTAATGTCAGTACTGTGCGTTTCAGAAAGCCCTGGA 758
OY      139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaLysGlnLys 158
        |||
DB      759 GTGGGATGTCATATATGCCATCAAGTGTGGCGGATGTCACAGGCCAGAGAGAGAG 818
OY      159 LeuLeuAlaGlnIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg 178
        |||
DB      819 CTGTGGCGGAGATCTCCAGTGATATCCAGCTCAATCCAGAGTCCGCTGACCTCCGT 878
OY      179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
        |||
DB      879 CAGGCCCTGGCAAAACATTGTATGACTCATATCAATAAATCCCTCCGCTGACCAAGGA 938
OY      198 LysAlaArgAlaIleLeuThrGlnLysThrThrAspLysSerProPheValIleTyrAsp 217
        |||
DB      939 AAGGGAGGAGCATCTTGACAGAGAAAGACACAGACAAATCACATTCCTTATCATGAC 998
OY      218 MetAsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGln 237
        |||
DB      999 ATGAATTCTCTTATGATGAGGAGAAATAAATCAAGTTCAAAACATCACCCCTGCGAG 1058
OY      238 GlnGlnSerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGln 257
        |||
DB      1059 GAGCAGAGCAAGAGAGTGGCCATCCGCACTTTCAGGGCTGCCAGTTTCGCTCCGTGGAG 1118
OY      258 AlaValGlnGlnIleThrGlnTyrAlaLysSerIleProGlnPheValAsnLeuAspLeu 277
        |||
DB      1119 GCTGTGCAAGAAATCATCAGAGATATGCCAAAGCATTTCCGTGTTTGTAAATTTTGACTTG 1178
OY      278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAla 297

```

```

DB      1179 AAGACACAGTAACCTCTCTCAAAATATGAGTCCACAGATCATTTACATGCTGGCC 1238
OY      298 SerLeuMetAsnLysAspLysValLeuIleSerGlnGlyGlnGlyPheMetThrArgGln 317
        |||
DB      1239 TCCTTGATGAATTAAGATGGGTTCTCATATATCCGAGGGCAAGGCTTCATGACAAGGGAG 1298
OY      318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGlnProLysPheGlnPheAla 337
        |||
DB      1299 TTCTTAAGAGCCTGGAAAGCCTTTTGGTGACTTTATGAGACCCAAAGTTTGAATTTGCT 1358
OY      338 ValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357
        |||
DB      1359 GTGAAGTTCAATGACACTGGAATTAATGATGACAGCAGCTTGGCAATFATTTATCTGTCATT 1418
OY      358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleLysAsp 377
        |||
DB      1419 ATTTCTAGTGGAGACCCGCCAGTTTGTGTAATGTGAACCCCATTTGAACATTCACAGAC 1478
OY      378 AsnLeuGlnAlaLeuGlnLeuGlnLeuLysLysAsnHisProGlnSerSerGlnLeu 397
        |||
DB      1479 AACCTGCTCAAGCCCTGAGCTCCAGCTGAAGCTGAACCATCCAGATCTTCACAGCTG 1538
OY      398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGln 417
        |||
DB      1539 TTTGCCAACCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTACGGAACACGCTGCAG 1598
OY      418 LeuLeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGln 437
        |||
DB      1599 CTACTGCAGGTATCAAGAAGGAGAGACAGACATGAGTTCACCCCTCTCTGCAGAG 1658
OY      438 IleTyrLysAspLeuTyr 443
        |||
DB      1659 ATCTCAAGACTTGTAC 1676

RESULT 4
US-09-128-142-1
; Sequence 1, Application US/09128142
; Patent No. 6294559
GENERAL INFORMATION:
APPLICANT: Smith, Roy G.
TITLE OF INVENTION: ANTIPROLIFERATIVE AGENTS ASSOCIATED WITH
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Jack L. Tribble
STREET: 126 E. Lincoln Ave., P.O. Box 2000
CITY: Rahway
STATE: New Jersey
COUNTRY: US
ZIP: 07065-0907
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: Power Macintosh 7500/100
SOFTWARE: Microsoft Word 6.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09-128-142
FILING DATE: 03-Aug-1998
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/844,007
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Tribble, Jack L.
REGISTRATION NUMBER: 32,633
REFERENCE/DOCKET NUMBER: <Unknown>
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 594-5321
TELEFAX: (908) 594-4720
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1434 base pairs

```

```

?      TYPE: nucleic acid
?      STRANDEDNESS: single
?
?      TOPOLOGY: linear
?
?      MOLECULE TYPE: DNA (genomic)
?      SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-128-142-1

```

Alignment Scores:	
Pred. No.:	2,88e+35
Score:	2089,00
Percent Similarity:	87,00%
Best Local Similarity:	87,00%
Query Match:	59,55%
DB:	4
	Gaps: 1

US-09-931-007A-1 (1-688) x US-09-128-142-1 (1-1434)

QY	29	MetHmethylAspThrGluMetProPheThrProHisnPhgIylIleSerSerVal	48
Db	1	ATGACCATGGTTGACACAGAGATCCATTCTGGGCCCAACAATTGGGATCAGCTCCGTG	60
QY	49	AspLeuSerValMetGluAspHisSerHisSerPheAspIleIysProPheThrVal	68
Db	61	GATCTCCCGGTAAGAAGAACCATCCACCTCCTTGATATCAAGCCCTCATACTAGTT	120
QY	69	AspPheSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAspPro	88
Db	121	GACTTCTCCACGATTTCTACTCCACATTCAGAAACATCTCCATTCCACAAAGACGATCCA	180
QY	89	ValValAlaAspTyrIlySTyrAspLeuLysLeuGlnIleTyrGlnSerAlaIleLysVal	108
Db	181	GTGGTTCCAGATTACAAAGTAGATGACCTTAACACTCAAAAGTACCAATGTCATCAAAAGTG	240
QY	109	GluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn-----	124
Db	241	GAGCTTCATCTCCACCTTATTATTCTGAGAAAGACTCAGCTCTCAATTAAGCCCATGA	300
QY	124	-----	124
Db	301	GAGCCTTCCAACTCCCTCATGCGCAATTGAATGTCTGTGTGAGATTAAGCTTCTGGA	360
QY	124	-----	124
Db	361	TTTCATATGAGATTCATGCTTGTGAAGATGCAAGGGTTCTTCGGAGAACATACGA	420
QY	125	-----ArgAsn	120
Db	421	TTGAGCTTATCTATGACAGATGTGATCTTAACGTGCGATCCCAAAAAAAGATGAAT	480
QY	127	LysCysGlnTyrCysArgPheGlnLysCysLeuAlaValAlaGlyMetSerHisAsnAlaIle	148
Db	481	AAATGTCAGTACTGTCCGTTTCAGAAATGCCCTTGCAGTGGGAGATGTCCTCAATAATGCCATC	540
QY	147	ArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSerAsp	166
Db	541	AGGTTTGGGCGGATGCCACAGGCCGGAAGAAAGAGAAAGCTGTGGCGGAGATCTCCACGTAT	600
QY	167	IleAspGlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAsp	186
Db	601	ATTCGACCAAGCTGAATCCAGAGATCCGCTGACCTCCGGGCCCTGGCAAAAACATTTGTTAAGAC	660
QY	187	SerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLys	206
Db	661	TCTATACATAAAGTCTTCCGCTGACCAAGCAAAAGCCGAGGATCTTGACACAGGAAG	720
QY	207	ThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetGlyLysAsp	226
Db	721	ACAACAGCAAAATACCATCTGTTATCTATGACATGATTCATTAAATGATGAGAGATAT	780
QY	227	LysIleLysPheLysHisIleThrProLeuGlnGlnGlnSerLysGluValAlaIleArg	246
Db	781	AAAATCAAGTTCAACACATATACCCCTCGAGGAGCAAGCAAAAGAGGTGGCCATCCGC	840

Qy	247	IllephglnlgyvsglnlphnleagserValglualValglnlIllethglnlTyrAla	266
Db	841	ATCTTTAGAGGGCTGCCAGTTTCCCTCCGCGAGGCTGTGCGAGAGATCACAGATTATGCC	900
Qy	267	LysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyr	286
Db	901	AAAGCATTCCTCGTTTTTGTAAATCTTGACTTGAGACGACCAAGTAACTCTCTCAAAATAT	960
Qy	287	GlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeu	306
Db	961	GGAATCCACAGATCATTTACAAAGCTCGGCCCTCTTGATCAATAAAGATGGGTTCTTC	1020
Qy	307	IleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPhe	326
Db	1021	ATATCCGAGGGCCCAAGGCTTCATGACCAAGGAGGTTTCTAAAGGCTCTCGAAAGCTTTT	1080
Qy	327	GlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeuAsp	346
Db	1081	GGTACTTTATGAGACCCCAAGTTTGAGTTTCTGCTGTGAGATTCAATGCACTGGAAATTAGAT	1140
Qy	347	AspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeu	366
Db	1141	GACAGCGACTGGCAATATTTATTTGCTGTGCATTATTCACAGTGGAGACCGCCGACGGTTTG	1200
Qy	367	LeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGln	386
Db	1201	CTGAAATCTGAAGCCCATTTGAAACATTTCAAGCAACCTGTCTCAAGGCCCTTGAGCTCCAG	1260
Qy	387	LeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThr	406
Db	1261	CTGAAGCTGAACCAACCTGTAGTCTCTACAGCTGTTTGGCCMAGCTGCTCAGAAATGACA	1320
Qy	407	AspLeuArgGlnIleValThrGluHisValGlnLeuLeuGlnValIleLysLysThrGlu	426
Db	1321	GACTCAGACAGATTTGTGCAGGACACGTCGACGACTGACGAGTATCAAAAGAGCGAG	1380
Qy	427	ThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr	443
Db	1381	ACAGACATGAGTCTTCACCCGCTCTGCAGAGAGATCTACAAAGACTTGTAC	1431
RESULT 5			
US-08-134-557D-1			
Sequence 1, Application US/08134557D			
Patent No. 6200802			
GENERAL INFORMATION:			
APPLICANT: Greene, Marianne E.			
APPLICANT: Blumberg, Bruce			
TITLE OF INVENTION: Human Peroxisome Proliferator Activated			
TITLE OF INVENTION: Receptor Gamma: Compositions and Methods			
NUMBER OF SEQUENCES: 8			
CORRESPONDENCE ADDRESS:			
ADDRESSEE: Rockey, Milnamow & Katz, Ltd.			
STREET: 2 Prudential Plaza, Suite 4700 180 N. Steeles			
CITY: Chicago			
STATE: IL			
COUNTRY: USA			
ZIP: 60601			
COMPUTER READABLE FORM:			
MEDIUM TYPE: Floppy disk			
COMPUTER: IBM PC compatible			
OPERATING SYSTEM: PC-DOS/MS-DOS			
SOFTWARE: PatentIn Release #1.0, Version #1.30			
CURRENT APPLICATION DATA:			
APPLICATION NUMBER: US/08/134,557D			
FILING DATE:			
CLASSIFICATION: 435			
ATTORNEY/AGENT INFORMATION:			
NAME: Katz, Martin L.			
REGISTRATION NUMBER: 25, 011			
REFERENCE/DOCKET NUMBER: ARCH:098			
TELECOMMUNICATION INFORMATION:			
TELEPHONE: 312-616-5400			
TELEFAX: 312-616-5460			

1: INFORMATION FOR SEQ ID NO: 1:  
2: SEQUENCE CHARACTERISTICS:  
3: LENGTH: 1844 base pairs  
4: TYPE: nucleic acid  
5: STRANDEDNESS: single  
6: TOPOLOGY: linear  
7: MOLECULE TYPE: DNA (genomic)  
8: FEATURE:  
9: NAME/KEY: CDS  
10: LOCATION: 179..1606  
11: US-08-134-557D-1  
  
Alignment Scores:  
Pred. No.: 2,46e-232 Length: 1844  
Score: 2065.50 Matches: 414  
Percent Similarity: 85.71% Conservative: 6  
Best Local Similarity: 84.49% Mismatches: 7  
Query Match: 58.88% Indels: 63  
DB: 4 Gaps: 2  
  
US-09-931-007A-1 (1-688) x US-08-134-557D-1 (1-1844)  
Oy 23 AlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrProThrAsn 42  
Db 155 GCCGCCGTGGCCGCGAATAATGACATGGTTGACACAGAGATCGATTCTGGCCACCAAC 214  
Oy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62  
Db 215 TTTGGATACAGCTCCGCGATCTCTCCGTAATGGAAAGCACTCCACCTCTGATATTC 274  
Oy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82  
Db 275 AAGCCCTTCACTACAGTGTGACTTCTCCAGCAATTTCTACTCCACATTTACAGAGACATTTCA 334  
Oy 83 PheThrArgThrAspProValValAlaAspTyrIleTyrAspIleLysLeuGlnGluTyr 102  
Db 335 TTCACAAACAAGATCCAGATGCTTGTCAGATTACAAAGTATGACCTGAACTTCAAGAGTAC 394  
Oy 103 GlnSerAlaIleLysValGluProAlaSerProTyrTyrSerGluLysThrGlnLeu 122  
Db 395 CAAAGTCGAATCAAAAGTGGAGCTGTCATCTCCACCTATTATTTCGAGAAGAGCTCAGCTC 454  
Oy 123 TyrAsn----- 124  
Db 455 TACATTAAGCTCTCATGAAGACCTTCCAACCTCCATGCGAATTGAATGTCGTCTGT 514  
Oy 124 ----- 124  
Db 515 GGAGATTAAAGCTTCTGATTTCTACTATGAGATTCATGCTTGTGAAGATGCAAGGTTTC 574  
Oy 124 ----- 124  
Db 575 TTCGGAGAACATACAGATGATAGCTTATGACAGATGATGATCTTAACGTGCGATC 634  
Oy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140  
Db 635 CACAAAAAAGTAAATAAATATGTCAGTACTGTGGTTTCAGAAATGCTTCGACGTGGG 694  
Oy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGlnLysLeuLeu 160  
Db 695 ATGTCCTAATATGCGATCAGGTTTGGGGGATCCACAGCGCCGAGAAAGGAAGCAACCTGTTG 754  
Oy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg---Ala 179  
Db 755 GCGAGATCTCCAGTGAATGACACAGCTGAATCCAGAGTCCGCTGACCTCCGTCAGGCC 814  
Oy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199  
Db 815 CTGCAAAACATTTGTATGACTCATATAAAGTCTTCCCGCTGACCAAGCAAGCGCG 874  
Oy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219  
Db 875 AGGCGATCTTGACAGGAAGACAACAGACAAATCACCATTCTGTTATCTATGACATGAAT 934

Oy 220 SerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGluGln 239  
Db 935 TCCTTAATGATGGGAGGAATTAATCAAGTTCAAAACATCACCCCCCTCAGAGAGAG 994  
Oy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259  
Db 995 AGCAAAAGAGTGGCCATCCGATCTTTCAGGGCTGCCAGTTTCGCTCGTGAGGCTGTG 1054  
Oy 260 GlnGluIleThrGlnGluTyrAlaLysSerIleProGlyPheValAsnLeuAsnAsp 279  
Db 1055 CAGGATCACAAGATATGCCCCAAAAGCATTCCTGTTTGTAAATCTTGACTGAAACGAC 1114  
Oy 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299  
Db 1115 CAAGTAACTCTCCCTCAATATGAGATCCACAGATCATTACACATCTGCTGCTGTG 1174  
Oy 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeu 319  
Db 1175 ATGAATTAAGATGGGGTCTCATATCCGAGGGCCAGAGCTTCATGACAAAGGAGTTTCTA 1234  
Oy 320 LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys 339  
Db 1235 AAGAGCTGCGAAAGCCTTTTGTGACTTATGAGAGCCCAAGTTGATTTGCTGTGAAG 1294  
Oy 340 PheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleLeu 359  
Db 1295 TTCATATGACATGGAATTATGATGACAGCGACTTGGCAATATTATTTCGTCTATTATCTC 1354  
Oy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379  
Db 1355 AGTGAGAGCCGCCAGGTTTCTGATGTGAAGCCCATTTGAAGCATTTCAAGAACCTTG 1414  
Oy 380 LeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPheAla 399  
Db 1415 CTACAGCCCTGGAGCTCCAGCTGAGCTGAACCACTGAGTCTCACACTGTTTGC 1474  
Oy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu 419  
Db 1475 AAGTGTCTCCAGAAATATGACAGACTGACAGATTTGTACAGGAACAGCTCAGCTACTG 1534  
Oy 420 GlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuGlnGluIleTyr 439  
Db 1535 CAGGTGATCAAGAAAGCAGACAGATGAGTCTTCCACCGCTCTGACAGAGATCTAC 1594  
Oy 440 LysAspLeuTyrAlaTyrPalaIleLeuThr 449  
Db 1595 AAGGACTTGTACTAGCAGAGAGTCTGAGC 1624  
  
RESULT 6  
US-09-255-392-1  
Sequence 1, Application US/09255392  
Patent No. 6214850  
GENERAL INFORMATION:  
APPLICANT: Evans, Ronald M.  
APPLICANT: Forman, Barry M.  
TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
STREET: 444 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: CA  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/255,392





NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
CITY: 444 South Flower Street, Suite 2000  
STATE: CA  
COUNTRY: USA  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/477,493  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Keller, Stephen E.  
REGISTRATION NUMBER: 31,192  
REFERENCE/DOCKET NUMBER: P41 9958  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 619-546-1995  
TELEFAX: 619-546-9392  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2005 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: both  
TOPOLOGY: both  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 352..1776  
US-08-477-493-1  
Alignment Scores:  
Pred. No.: 3,21e-232 Length: 2005  
Score: 2065.00 Matches: 408  
Percent Similarity: 86.40% Conservative: 5  
Best Local Similarity: 85.36% Mismatches: 3  
Query Match: 58.87% Indels: 62  
Gaps: 1  
US-09-931-007a-1 (1-688) x US-08-477-493-1 (1-2005)  
QY 28 GluMetThrMetValAspThrGluMetProPheThrAspPheGlyLeuSerSer 47  
Db 343 GAAATTACCATGCTGTTGACACACAGATGCCATTCTGCCCCACCACTTCGGAAATCAGCTCT 402  
QY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrThr 67  
Db 403 GTGGACCTCTCCGTGATGGAAGACACACGCGATTCCTTGACATCAACGCCCTTACCA 462  
QY 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAsp 87  
Db 463 GTTGAATTTCTCCAGCATTTCTGCTCCACACATGAAGACATTCCTCAACAAGACCTGAC 522  
QY 88 ProValValAlaAspTyrLysTyrAspLeuLysLeuGlnIuTyrGlnSerAlaIleLys 107  
Db 523 CCAATGGTGCTGATTACAAATATGACCTGAAGCTCCAAGAATTACCAAGAAGTGCATCAA 582  
QY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124  
Db 583 GTAGAACCCTGCATCTCCACCTTATATCTGAAAGACCCAGGCTGTACAAACAGCGCTCAT 642  
QY 124 ----- 124  
Db 643 GAAGAACCTTCTAACTCCCTCATGCGCATGTAGTCCGAGTCTGTGGGATATAAGCATCA 702  
QY 124 ----- 124  
Db 703 GGCTTCCACTATGAGTTCATGCTTGTGAAGAGATGCAAGGCTTTTTCGCAAGAACCATC 762

QY 125 -----Arg 125  
Db 763 CGATTGAACCTTATTATGATAGTGTGATCTTAACCTCCGGATCCACAAAAAAGTGA 822  
QY 126 AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145  
Db 823 AATAATGTAGTACTGTGCTGAGAAAGTCCCTTGCTGTGGGATGCTTCACAAATGCC 882  
QY 146 IleArgPheGlyArgMetProGlnIaGluLysGluLysLeuLeuAlaGluIleSerSer 165  
Db 883 ATCAGGTTTGGGGGATGCCACAGGCCGAGAAGGAGAGCTTTGGCCGAGATCTCCAGT 942  
QY 166 AspIleAspGlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyr 185  
Db 943 GATATCGACACAGCTGAACCCAGACGCTCTGATCTGCCAGCCCTTGCAAAAGCATTTGTAT 1002  
QY 186 AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205  
Db 1003 GACTCATACATAAAGTCCCTCCCGCTGACCAAGCCAGGCGAGGCGATCTTGACAGA 1062  
QY 206 LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetMetGlyGlu 225  
Db 1063 AAGACACGGACAAATVCAACCATTTGTCTATCAGCAGCATGAATTCCTTAATGATGGAGAA 1122  
QY 226 AspLysIleLysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIle 245  
Db 1123 GATAAATCAAGTTCAACCATATACACCCCTCGCAGGAGCAGAGCAAGAGTGCGCCATC 1182  
QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265  
Db 1183 CGAATTTTCAAGGGGTGCAGATTGCGATCCGTAGAAAGCCGTGCAAGAGATCACAGATAT 1242  
QY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285  
Db 1243 GCCAAAATATCCCTGGTTTATTACCTGATTGAATGACCAAGTACCTGCTGCACAG 1302  
QY 286 TyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305  
Db 1303 TATGGTGTCCATGAGATATATACACGATCTCGGCTCCCTGATGAAATAAAGATGAGATC 1362  
QY 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325  
Db 1363 CTCATCTCAGAGGGCCAGAGATTCATGACACAGGAGGTTCTCAAAAGCCTCGGAAGCCC 1422  
QY 326 PheGlyAspPheMetGluProLysPheGlnPheAlaValLysPheAsnAlaLeuGlnLeu 345  
Db 1423 TTGGTGACTTATGAGACCTTAAGTTTGAATTTGCTGTGAAGTTCAATGCACTGGAAATTA 1482  
QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365  
Db 1483 GATGACATGACTTGGCATATTTATATACCTGTCATTTATCTCAGTGGAGACCGCCAGGC 1542  
QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeu 385  
Db 1543 TTGGTGAACGTGAAGCCCATGAGAGACATCCAAAGACACCTGTCGAGGCGCTGGAACCTG 1602  
QY 386 GlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMet 405  
Db 1603 CAGCTCAAGCTGAATCACCCAGAGCTCTCTCAGCTGTTCGCAAGAGTCTCCAAAGAATG 1662  
QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeuGlnValIleLysLysThr 425  
Db 1663 ACAGACCTCAGCAGAGATGCTGACAGAGACGCTGACGCTACTGTCATGTGATCAAGAAGCA 1722  
QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443  
Db 1723 GAGACAGACATGAGCCTTACCCCTGCTCCAGAGATCTCAAGAGACTGTAT 1776  
RESULT 9  
US-08-465-375-1  
Sequence 1, Application US/08465375A  
Patent No. 6022897







```
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1776)
US-09-788-070-1

Alignment Scores:
Pred. No.: 3, 21e-232 Length: 2005
Score: 2065.00 Matches: 408
Percent Similarity: 86.40% Conservative: 5
Best Local Similarity: 85.36% Mismatches: 3
Query Match: 58.87% Indels: 62
DB: Gaps: 1

US-09-931-007a-1 (1-688) x US-09-788-070-1 (1-2005)

OY 28 GUMetThmEValAspHrGluMetProPhetRPProThraSnbheGlyIleSerSer 47
    |||:::|||||
DB 343 GAATTTACCATGGTTGACACAGAGATGCCATTCGGCCCCCAACTTCGGAATCAGCTCT 402
OY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPhetRThr 67
    |||:::|||||
DB 403 GTGGACCTCTCCGTGATGGAAGACCACTGCCATTCCTTTGACATCAAGCCCTTTACCA 462
OY 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPhetRThrAsp 87
    |||:::|||||
DB 463 GTGATTTTCCAGCATTTCTGCTCCACACTATGAAGACATTCATTCACAAGAGCTGAC 522
OY 88 ProValValAlaAspTyrIleTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLys 107
    |||:::|||||
DB 523 CCATGGTGGTGATTAACAAATATGACCTGAAGCTCCAAAGAAATACCAAGTGGCATCAA 582
OY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
    |||:::|||||
DB 583 GTGAACCTGCATCTCCACTTATTATTCTGAAGAAGACCAAGCTCTACACAGAGCTCAT 642
OY 124 ----- 124
DB 643 GAAGACCTTCTAATCCCTCATGGCCATGATGCCGAGTGTGTGGGGATTAAGCATCA 702
OY 124 ----- 124
DB 703 GGGTTCACATGAGATTCAGTCTGTGAAGATGCAAGGTTTTTCCGAGAACCATC 762
OY 125 -----Arg 125
DB 763 CGATTACACCTTATTATATAGTGTGATCTTAAGTCCGGGATCCACAAAAAAGTAGA 822
OY 126 AsnLysCysGlnTyrIleCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
    |||:::|||||
DB 823 AATTAATGTGAGTACTGTGGTTTCAGAAAGTGCTGTGGTGGCATGTCTCATCATGCC 882
OY 146 IleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSer 165
    |||:::|||||
DB 883 ATCAGGTTTGGGGGATGGCAGACAGGCCGGAAGGAAGAAAGCTGTGGCGAGATCTCCAG 942
OY 166 AspIleAspGlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyr 185
    |||:::|||||
DB 943 GATATGACACCTGACAGCAGAGTGTGCTGATCTCGAGCCCTGGCAAGCATTTGTAT 1002
OY 186 AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
    |||:::|||||
DB 1003 GACTCATACATTAAGTCTCCCGCTGACCAAGCAAGCGGCGAGTCTTGACAGAGA 1062
OY 206 LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetLeuTyrGlu 225
    |||:::|||||
DB 1063 AAGACACGACGAAATACCATTTTGCATCTAGACATGAATTCCTTAATGATGGAGAA 1122
OY 226 AspLysIleLysPheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIle 245
    |||:::|||||
DB 1123 GATTAATAATCAAGTCAAAATATATACCCCCCTGCGAGAGCAGAGCAAAAGGTGGCATC 1182
```

```
OY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265
    |||:::|||||
DB 1183 CGAATTTTCAAGGATGCCAGTTTCATCCGTGAAGCCGTGCACAGATCAACAGATAT 1242
OY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
    |||:::|||||
DB 1243 GCCAAATATCCCTGGTTTCATTAACCTTGATTTGAATACCAAGGAGACTGCTCCAG 1302
OY 286 TyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetLysAspGlyVal 305
    |||:::|||||
DB 1303 TATGTGTCCATAGATCATCTACAGATGCTGGCTCCCTGATGAATTAAGATGGAGTC 1362
OY 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuAlyLysPro 325
    |||:::|||||
DB 1363 CTCATCTCAGAGGCCAAGGATTCATGACAGGAGTCTCTCAAAAGCCCTGGCAAGCCC 1422
OY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
    |||:::|||||
DB 1423 TTTGGTACTTGTATGAGAGCTTAAGTTTGAGTTTGTGTGAAGTTCAATGCATCGAATTA 1482
OY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
    |||:::|||||
DB 1483 GATGACAGTACTGGCTATATTATAGCTGTCTATTCTCATTTCTGAGAGACGCCCAAGC 1542
OY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385
    |||:::|||||
DB 1543 TTCTGACACCTGAAGCCCATCGAGGACATCCAGACACACACTGCTGCAAGGCCCTGGAAC 1602
OY 386 GlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPheAlaLysLeuLeuGlnLysMet 405
    |||:::|||||
DB 1603 CACCTGAGGTGAATCACCCAGAGTCTCTCACCTGTCCCAAGGTGCTCCAGAAAGATG 1662
OY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeuGlnValIleLysLysThr 425
    |||:::|||||
DB 1663 ACAGACCTCAGGACAGATCTGTCACAGACAGCAGCTGACACTGATGATCAAGAGACA 1722
OY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
    |||:::|||||
DB 1723 GAGACAGACATGAGCCTTCACCCCTGCTCCAGAGATCTACAAAGGACTGTGAT 1776

RESULT 11
US-08-484-200-3
; Sequence 3, Application US/08484200
; Patent No. 5861274
; GENERAL INFORMATION:
; APPLICANT: EVANS, RONALD M.
; APPLICANT: FORMAN, BARRY M.
; APPLICANT: KLEWER, STEVEN A.
; APPLICANT: ONG, ESTELITA S.
; TITLE OF INVENTION: NOVEL MEMBERS OF THE STEROID/THYROID
; TITLE OF INVENTION: SUPERFAMILY AND USES THEREFOR
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK
; STREET: 444 SOUTH FLOWER STREET, SUITE 2000
; CITY: LOS ANGELES
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,200
; FILING DATE: 07-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: REITER, STEPHEN E.
; REGISTRATION NUMBER: 31,192
; REFERENCE NUMBER: P41 9971
; TELECOMMUNICATION INFORMATION:
```

```

? TELEPHONE: 619-546-1995
? TELFAX: 619-546-9392
? INFORMATION FOR SEQ ID NO:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2012 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: both
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 263..1582
US-08-484-200-3

```

Alignment Scores:	
Pred. No.:	2,83e-116
Score:	1076.50
Percent Similarity:	80.06%
Best local Similarity:	62.27%
Query Match:	30.69%
DB:	2
Length:	2012
Matches:	58
Conservative:	203
Mismatches:	64
Indels:	1
Gaps:	1

US-09-931-007A-1 (1-688) x US-08-484-200-3 (1-2012)

Oy	119	LysThrGlnLeuIleuYrAsnAlaGAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla	138
Db	605	AAGATCCAGAAAGAAAGCCCAACAGTGTCACTACTGCCGCTTCCAGAGTGCCTGGCA	664
Oy	139	ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnInlaGluLysGluLys	158
Db	665	CTCGGCATGTGCGCAACAGCGATTCGGTTTGGAGGATGTGGGAGGCGAGAAAGAGAG	724
Oy	159	LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeu	177
Db	725	CTGGTGGGGGGGCTGACGTGGCAGCGAGGGGTGGCAGCACAACCCCAAGCTGGCGGAC	784
Oy	178	ArgAlaLeuAlaLysHisLeuIleuYrAspSerTyrIleLysSerPheProIleuYrAspAla	197
Db	785	AAGGCTTTCTTAAGCAACATCTTACACGGCTACTCTGAAATACTTCAACATGTCCAAAGA	844
Oy	198	LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp	217
Db	845	AAGCGCCGAGACATCTTCACCGGGAAGTCCAGGCACACAGCAACCTTTGTTCATCCAGC	904
Oy	218	MetLysSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGln	237
Db	905	ATTCGAGACACTGTGGCCAGCGAGAAAGGGCTGTGTGMAACGTGTGTGAACGGGGCTG	964
Oy	238	GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu	257
Db	965	CCGCGCTTACACAGGATTCAGTGTGCACGTGTTTACCGCTGCCAGTCCACCAAGGGAG	1022
Oy	258	AlaValGlnGluIleThrGlyTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu	277
Db	1025	ACAATCCGAGAGGTCACCGAGTTCGCCAABAATCCCAACTTTCAGACGCTCTTCTTC	1084
Oy	278	AsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAla	297
Db	1085	AATGACCAGGAGACCTCTTCACAGTATGCGGTGCACAGAGCCATCTTCCATGCTGCC	1144
Oy	298	SerLeuMetLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGlu	317
Db	1145	TTCATCTCTCAACAAAGACGGGGCTGTGGTGGCCAAAGCAGTGGCTCTCCACCCACGAG	1204
Oy	318	PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla	337
Db	1205	TTCCTTGGGAAGTCTCCGCCAAGCCCTTCAGTACATATTTAGAGCCCAAGTTCCAGTTTGC	1264
Oy	338	ValLysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIle	357
Db	1265	GTCAGATTTCATGTGGCTGGAGCTGATGACAGTACGACCTGGCGCTTTCATCCGGGGC	1322
Oy	358	IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluLysPheGlnAsp	377

```

Db 1325 ATTCGTGTGGAGACCGGCGCCAGCTCATGAATGTGCCCCAGGTGAAGACCATCCAGGAC 1384
Oy 378 AsnLeuLeuGlnAlaLeuGlnGluLeuGlnLeuLysLeuAsnHisProGlnLysSerSerLeuLeu 397
Db 1385 ACCATTCTGGGGGCTCTAAGATTCCATCTGCAGCTCAACCCAGCTGACGACCGCATGACTTC 1444
Oy 398 PheAlaLysLeuLeuGlnLysMetLtrAspLeuArgGlnIleValThrGlnHisValGln 417
Db 1445 TTCCCCAAGCTGCTGCAGAGATGGCAGACCTCGGCGACCTGGTCTCATGAGCATCCGAC 1504
Oy 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGln 437
Db 1505 ATGATGACAGTGGCTAAAGAGAGACGAGAGTGAACCTTGCTGACCCCTGCTCCAGGAA 1564
Oy 438 IleTyrLysAspLeuTyr 443
Db 1565 ATCTACAAGGACATGTAC 1582

RESULT 12
US-08-333-358-9
Sequence 9, Application US/08333358
Patent No. 5571696
GENERAL INFORMATION:
APPLICANT: EVANS Ph.D., RONALD M.
APPLICANT: MANGELSDORF Ph.D., DAVID J.
APPLICANT: ONG Ms., ESTELITA S.
APPLICANT: ORO Ph.D., ANTHONY E.
APPLICANT: BORMEYER Ph.D., UWE K.
APPLICANT: GIGUERE Ph.D., VINCENT NMN
APPLICANT: YAO Mr., TSO-PANG NMN
TITLE OF INVENTION: NOVEL RECEPTORS
NUMBER OF SEQUENCES: 14
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark
STREET: 444 So. Flower St., Suite 2000
CITY: Los Angeles
STATE: CA
COUNTRY: US
ZIP: 90071-2921
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/333,358
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/761,068
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Reiter Ph.D., Stephen E.
REGISTRATION NUMBER: 31192
REFERENCE/DOCKET NUMBER: P31 8936
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 2009 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
IMMEDIATE SOURCE:
CLONE: XR4 (XR4.SEG)
FEATURE:
NAME/KEY: CDS
LOCATION: 263..1582
US-08-333-358-9

```



```

Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluysGluys 158
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 665 CTCGGCATGTCACACAAAGCGCTTTCGACGATGCCGAGCGGACGAGAAGAGGAG 724
Oy 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeu 177
    ::::::::::::::::::::
Db 725 CTGGGGCGGGGCTACTGCTCCAGAGGGGCTCCAGACACCCCGCTGACGCTGAC 784
Oy 178 ArgAlaLeuAlaHisHisLeuTyraSerTyriLeysSerPheProLeuThrIysAla 197
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 785 AAGGCTTCTCTACACACATCTACACAGCCTTACCTGAAAACCTTCAACATGACCAAAAG 844
Oy 198 IysAlaArgAlaIleLeuThrGlyIysThrThrAspLysSerProPheValIleTyraSp 217
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 845 AAGGCCCGGAGATCCTTCCACGGCAAGTCCACCAACGACGACCCCTTTCATCCAGCAG 904
Oy 218 MetAsnSerLeuMetMetCylGluAspLysIleLysPheLysHisIleThrProLeuGln 237
    ::::::::::::::::::::
Db 905 ATCGAGACACTGTGGCAGGACGAGAAAGGCGCTGTGTGAAACAGCTGTGACGCGCG 964
Oy 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 965 CCTTACAAAC---GAGATCAGTGTGACAGCTTCTTACCGCTGCCAGTCCACCAAGTGGAG 1021
Oy 258 AlaValGlnGluIleThrGluTyraLysSerIleProGlyPheValAsnLeuAspLeu 277
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1022 ACAGTCCGAGACCTACCGACGATTCGCGCAAGACATCCCACTTGCAGCGCTTCTCTC 1081
Oy 278 AsnAspGlnValThrLeuLeuLysTyriGlyValHisGluIleIleTyriThrMetLeuAla 297
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1082 AATGACACAGGTACCTCTCAAGATGCGGTGACAGAGGACCTTTCCTCATCTGCTGGCC 1141
Oy 298 SerLeuMetAsnLysAspLysValLeuIleSerGluGlyPheMetThrArgGlu 317
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1142 TCCATGCTCAACAAAGACGGCGTGTGCGGACGACGAGTGGCTTCTTCCACCGCAG 1201
Oy 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1202 TTCTTGGCAAGTCTCCGCAAGCCCTTCACTAGTACATCATGAGCCCAAGTTTCGATTGCT 1261
Oy 338 ValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1262 GTCAAGTTTAAAGCCGTGAGTCTGATGACAGTACGCTGCGCTTCTTCAATCGCGGCATC 1321
Oy 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1322 ATTCGTGTGAGAGACCGCGGACGCTCATGATGTGCCAGGTAGAACCATCCAGGAC 1381
Oy 378 AsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeu 397
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1382 ACCATTCCTCGGGCTCTAGAAATTCATCTGCAAGTCAACCCCTGACAGCCAGTACCTC 1441
Oy 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1442 TTCCCCAACGCTCTCAGAGATGCGACCTGCGGACGCTGTGCTCATGAGATGCGCCAG 1501
Oy 418 LeuLeuGlnValIleIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1502 ATGATGCACTGCTAAGACAGACGAGAGTGAACCTTCTGACACCCCTGCTCCAGGA 1561
Oy 438 IleTyriLysAspLeuTyri 443
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 1562 ATCTCAAGACATGTGAC 1579

```

```

? APPLICANT: BORGMEYER Ph.D., UWE K.
? APPLICANT: GIGUERE Ph.D., VINCENT NMN
? APPLICANT: YAO M., TSO-PANG NMN
? TITLE OF INVENTION: NOVEL RECEPTORS
? NUMBER OF SEQUENCES: 14
? CORRESPONDENCE ADDRESS:
? ADDRESS: Pretty, Schroeder, Brueggemann & Clark
? STREET: 444 So. Flower St., Suite 2000
? CITY: Los Angeles
? STATE: CA
? COUNTRY: US
? ZIP: 90071-2921
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: PatentIn Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/694,501
? FILING DATE: 07-AUG-1996
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US/08/333,358
? FILING DATE:
? APPLICATION NUMBER: US/07/761,068
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: Reiter Ph.D., Stephen E.
? REGISTRATION NUMBER: 31192
? REFERENCE/DOCKET NUMBER: P31 8936
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (619) 535-9001
? TELEFAX: (619) 535-8949
? INFORMATION FOR SEQ ID NO: 9:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 2009 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: cDNA
? IMMEDIATE SOURCE:
? CLONE: XR4 (XR4.SBG)
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 263..1582
? US-08-694-501-9
?
? Alignment Scores:
? Pred. No.: 1,256-115 Length: 2009
? Score: 1071.00 Matches: 202
? Percent Similarity: 80.37% Conservative: 60
? Best Local Similarity: 61.96% Mismatches: 62
? Query Match: 30.53% Indels: 2
? DB: 1 Gaps: 2
?
? US-09-931-007a-1 (1-688) x US-08-694-501-9 (1-2009)
Oy 119 LysThrGlnLeuTyraAsnArgAsnLysCysGlnTyriGlyArgPheGlnLysCysLeuAla 138
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 605 AAGATCCACAGAGAGAACCGCAACAAGTGTCACTGCTGCTTCCAGAAAGTCCCTGGCA 664
Oy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluysGluys 158
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 665 CTCGGCATGTCACACACGCTTCCGCTTTCGACGATGCCGAGCGGACGAGAAGAGGAG 724
Oy 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeu 177
    ::::::::::::::::::::
Db 725 CTGGGGCGGGGCTACTGCTCCAGAGGGGCTCCAGACACCCCGCTGACGCTGAC 784
Oy 178 ArgAlaLeuAlaHisHisLeuTyraSerTyriLeysSerPheProLeuThrIysAla 197
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
Db 785 AAGGCTTCTCTACACACATCTACACAGCCTTACCTGAAAACCTTCAACATGACCAAAAG 844

```



Db	526	ATGCCAAGATCTGGAGAAAGCAAAACTGAAAGCAAGAAATCTTCACTGTGCACATGACATA	585
Oy	170	LeuasnProgluSerAlaaspLeuArqAlaLeuAlaLysHisLeuTyAspSerTyIle	189
Db	586	GAACATCTTGCAAACTGCAGATCTCAAAATCTCTGGCCAAAGAAATCTACAGAGCCATTC	645
Oy	190	LysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrasp	209
Db	646	AAGAACTTCAACATGACAAAGAGTCGCAAGCCGGGGTCATCTCTTCAGAAAGGCCAGTAC	705
Oy	210	LysSerProPheValIleTyAspMetAsnSerLeuMetMetGlyLysAspLysIleLys	229
Db	706	AATCACCTTTTGTGCATACATGATATGAGAGACACGCTGTATGCTGTAGAGACGCTGTGTG	765
Oy	230	PheLysHisIleThrProLeuGlnGlnSerLysGluValAlaIleArgIlePheGln	249
Db	766	GCCAAAG---CTGGTGCCCAATGGCATTCGCAACCAAGAGGCGGAGGCTGGCATCTTTCAC	822
Oy	250	GlyLysGlnPheArgSerValGluAlaValGlnGluIleThrGlyTyAlaLysSerIle	269
Db	823	TCGTGCCAGTGCACGTACGTGGTGACCGTCACCGAGCTACCGGAATTCGCCAAGCCATC	882
Oy	270	ProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyGlyValHis	289
Db	883	CCAGGCTTCGCAAACTTGGAGCAGTCAAGATGACATTCCTCAAAATACGAGATTAT	942
Oy	290	GluIleIleTyThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlu	309
Db	943	GAGGCCATATTCGCCATGCTGCTCTTGATGATCAACAAAGAGGAGTCTGTATGCGGAT	1002
Oy	310	GlyGlnGlyPheMetThrArgGlnPheLeuLysSerLeuArgLysProPheGlyAspPhe	329
Db	1003	GCAAAATGGGTTTAACTCGTGAATTCCTAAAAAGCCTAAGAAACGTTCTGTGATATC	1062
Oy	330	MetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGlnLeuAspSerAsp	349
Db	1063	ATGGAACCCAAAGTTGATTTGGCATGAAGATTCATAGCACTGGAACTGGAGACAGTAT	1122
Oy	350	LeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnVal	369
Db	1123	ATTCGCCCTTTTGGCTGCTCATCATTTTGTGTGGAGATCGTCTCGCCCTTCTAATACGA	1182
Oy	370	LysPheIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeu	389
Db	1183	GGACACATTTGAAAAAATGAGGAGGATGTATGTAACATGTCTCAGACCTCACCTGACGAGC	1242
Oy	390	AsnHisProgluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArg	409
Db	1243	AACCAACCCGACGATATCTTCTCTTCCCAAAACTTCTTCAAAAAATGACAGACCTCGG	1302
Oy	410	GlnIleValThrGlnHisValGlnLeuLeuGlnValIleLysLysThrGlyThrAspMet	429
Db	1303	CAGCTGGAGAGGAGCATCGCCAGCGTGTGCGAGATCATCAAAACAGCAGCATCGCATCT	1362
Oy	430	SerLeuHisProLeuLeuGlnGlnIleTyLysAspLeuTy	443
Db	1363	GGCGTTCACCCCGCTACTGCGAGGATCTACAGGAGCATGTAC	1404

Search completed: February 25, 2003, 04:10:34  
Job time : 87 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: February 25, 2003, 02:38:25 ; Search time 61 Seconds

(without alignments)  
1502.892 Million cell updates/sec

Title: US-09-931-007a-1

Perfect score: 3508  
Sequence: 1 MERTLGSDPIPESDSFTDT.....KTEIDMSLHPLLQETIKDLY 688Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq.101002:\*

1: /SID2/gcgdata/geneseq/genesep-emb1/AA1980.DAT:\*  
2: /SID2/gcgdata/geneseq/genesep-emb1/AA1981.DAT:\*  
3: /SID2/gcgdata/geneseq/genesep-emb1/AA1982.DAT:\*  
4: /SID2/gcgdata/geneseq/genesep-emb1/AA1983.DAT:\*  
5: /SID2/gcgdata/geneseq/genesep-emb1/AA1984.DAT:\*  
6: /SID2/gcgdata/geneseq/genesep-emb1/AA1985.DAT:\*  
7: /SID2/gcgdata/geneseq/genesep-emb1/AA1986.DAT:\*  
8: /SID2/gcgdata/geneseq/genesep-emb1/AA1987.DAT:\*  
9: /SID2/gcgdata/geneseq/genesep-emb1/AA1988.DAT:\*  
10: /SID2/gcgdata/geneseq/genesep-emb1/AA1989.DAT:\*  
11: /SID2/gcgdata/geneseq/genesep-emb1/AA1990.DAT:\*  
12: /SID2/gcgdata/geneseq/genesep-emb1/AA1991.DAT:\*  
13: /SID2/gcgdata/geneseq/genesep-emb1/AA1992.DAT:\*  
14: /SID2/gcgdata/geneseq/genesep-emb1/AA1993.DAT:\*  
15: /SID2/gcgdata/geneseq/genesep-emb1/AA1994.DAT:\*  
16: /SID2/gcgdata/geneseq/genesep-emb1/AA1995.DAT:\*  
17: /SID2/gcgdata/geneseq/genesep-emb1/AA1996.DAT:\*  
18: /SID2/gcgdata/geneseq/genesep-emb1/AA1997.DAT:\*  
19: /SID2/gcgdata/geneseq/genesep-emb1/AA1998.DAT:\*  
20: /SID2/gcgdata/geneseq/genesep-emb1/AA1999.DAT:\*  
21: /SID2/gcgdata/geneseq/genesep-emb1/AA2000.DAT:\*  
22: /SID2/gcgdata/geneseq/genesep-emb1/AA2001.DAT:\*  
23: /SID2/gcgdata/geneseq/genesep-emb1/AA2002.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3467	98.8	750	22	Human PPARgamma2ga
2	3467	98.8	750	23	Human peroxisome p
3	2351	67.0	467	22	Human PPARgamma2ga
4	2231	63.6	505	17	Peroxisome Prolife
5	2231	63.6	505	22	Human peroxisome p
6	2231	63.6	505	22	Human PPARgamma pr
7	2231	63.6	505	22	Human PPARgamma
8	2106.5	60.0	811	22	Human PAX8e3(-exon
9	2106.5	60.0	874	22	Human PAX8e3-PPARg
10	2099	59.8	777	22	Human PAX8e7-PPARg

11	2094	59.7	494	17	AA99323
12	2094	59.7	840	22	AA85794
13	2089	59.5	477	17	AA99324
14	2089	59.5	477	17	AA99327
15	2089	59.5	477	17	AAE1867
16	2079	59.3	475	17	AA99325
17	2079	59.3	475	17	AA99328
18	2054.5	58.6	478	20	AA905471
19	2054.5	58.6	478	22	AA85802
20	2054.5	58.6	478	22	AA98350
21	2054.5	58.6	478	22	AAE00912
22	2054	58.6	475	17	AA905777
23	2054	58.6	475	17	AA92478
24	2054	58.6	475	20	AAW9596
25	2044.5	58.3	476	22	AAW67656
26	1961	55.9	475	20	AA921635
27	1569	44.7	313	22	AA674068
28	1108.5	31.6	420	20	AA921634
29	1101.5	31.4	441	14	AA841875
30	1101.5	31.4	441	17	AA898214
31	1101.5	31.4	441	20	AA905472
32	1076.5	30.7	440	17	AA92479
33	1076.5	30.7	440	20	AAW9597
34	1076.5	30.7	468	16	AA874053
35	1076.5	30.7	468	22	AA820342
36	1074.5	30.6	468	20	AA921633
37	1074.5	30.6	468	20	AA905470
38	1056	30.1	439	14	AA833745
39	524	14.9	121	21	AA853489
40	512.5	14.6	128	21	AA853442
41	404.5	11.5	579	16	AA871565
42	392	11.2	2065	22	AA863705
43	383.5	10.9	1237	12	AA813791
44	343	9.8	777	22	AA81894
45	342.5	9.8	548	14	AA835741

## ALIGNMENTS

RESULT 1	
AA876987	
ID	AA876987 standard; Protein: 750 AA.
XX	
AC	AA876987;
XX	
DT	22-JUL-2002 (first entry)
XX	
DE	Human PPARgamma2gamma2.
XX	
KW	PPAR response element; PPAR: vaccine; gene therapy; human;
KW	peroxisome proliferator-activated receptor.
OS	Homo sapiens.
XX	
PN	WO20078986-A1.
XX	
PD	28-DEC-2000.
XX	
PF	22-JUN-2000; 2000WO-FR01744.
XX	
PR	22-JUN-1999; 99FR-0007957.
XX	
PR	20-AUG-1999; 99US-0149721.
XX	
PA	(AVENT ) AVENTIS PHARMA SA.
XX	
PI	Datell R, Crouzet J, Steals B, Mahfoudi A;
XX	
DR	WPI: 2001-091574/10.
XX	
PT	Composition providing inducible expression of a nucleic acid, useful in
PT	gene therapy, uses minimal promoter with peroxisome
PT	proliferator-activated receptor response elements





```

Db      1 MGETLADSPIDRESDFTLNANISQEMTMVDEMPFMTNNGISSVDLSVMEHSHS 60
QY      61 DIKFTTVOFSSTPHYEDIPFTRTDPVADYKYLKLOEYSAIKVEPASPPYSEKT 120
      61 DIKFTTVOFSSTPHYEDIPFTRTDPVADYKYLKLOEYSAIKVEPASPPYSEKT 120
QY      121 QLYN----- 124
      121 QLYN----- 124
Db      121 QLYNKHHEPSNLSMAIECVGCDKASGFHYVHACEGCKGFFRRITRLKLIYRCDLNC 180
QY      125 -----RNKCQYCRFQKCLAVGSHNAIRFGMPQAEKEKLAIEISSDIDQNPESADLR 178
      181 RIHKSRNKCQYCRFQKCLAVGSHNAIRFGMPQAEKEKLAIEISSDIDQNPESADLR 240
QY      179 ALAKHLYDSYIKSEPLTKAKARAILTGKTTDKSPFIYDMNSLMGEDKIKFHNITPLOB 238
      241 ALAKHLYDSYIKSEPLTKAKARAILTGKTTDKSPFIYDMNSLMGEDKIKFHNITPLOB 300
QY      239 QSEKVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQVTLTKYGVHEIITMLAS 298
      301 QSEKVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQVTLTKYGVHEIITMLAS 360
QY      299 LMKKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFNALDELDDSLAIFIAVII 358
      361 LMKKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFNALDELDDSLAIFIAVII 420
QY      359 LSGDRGLNVKPIEDIONLLOALELOKLNHPRESSOLFATLQKMTDLRQIVTEHVOL 418
      421 LSGDRGLNVKPIEDIONLLOALELOKLNHPRESSOLFATLQKMTDLRQIVTEHVOL 480
QY      419 LOYIKKETDMSLHPLQEIYKDYAMAILTGKTTDKSPFIYDMNSLMGEDKIKFKHI 478
      481 LOYIKKETDMSLHPLQEIYKDYAMAILTGKTTDKSPFIYDMNSLMGEDKIKFKHI 540
QY      479 TPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQVTLTKYGVHEIIT 538
      541 TPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQVTLTKYGVHEIIT 600
QY      539 TMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFNALDELDDSLAIF 598
      601 TMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFNALDELDDSLAIF 660
QY      599 IAVIILSGDRPGLNVKPIEDIONLLOALELOKLNHPRESSOLFATLQKMTDLRQIVT 658
      661 IAVIILSGDRPGLNVKPIEDIONLLOALELOKLNHPRESSOLFATLQKMTDLRQIVT 720
QY      659 EHVOLLQVIAKKTETDMSLHPLQEIYKDY 688
      721 EHVOLLQVIAKKTETDMSLHPLQEIYKDY 750
Db

```

## RESULT 3

```

ABB76988
ID      ABB76988 standard; Protein: 467 AA.
XX
AC      ABB76988;
XX
XX      22-JUL-2002 (first entry)
DT
XX
DE      Human PPARgamma2gamma2 E/F domain.
KW      PPAR response element; PPAR: vaccine; gene therapy; human;
KM      peroxisome proliferator-activated receptor.
XX
OS      Homo sapiens.
XX
PN      WO200078986-A1.
XX
PD      28-DEC-2000.
XX
PF      22-JUN-2000; 2000WO-FR01744.
XX

```

```

PR      22-JUN-1999; 99FR-0007957.
PR      20-AUG-1999; 99US-0149721.
XX
XX      (AVET ) AVENTIS PHARMA SA.
XX
PI      Dartell R, Crouzet J, Staels B, Mahfoudi A;
XX
DR      WPI: 2001-091574/10.
XX
XX
PT      Composition providing inducible expression of a nucleic acid, useful in
PT      gene therapy, uses minimal promoter with peroxisome
PT      proliferator-activated receptor response elements
PS
XX
XX      Example 5; Page 34; 94pp; French.
XX
CC      The present invention relates to a composition (A) comprising a component
CC      (A1) containing a nucleic acid (1) controlled by an inducible promoter
CC      that consists of a PPAR (peroxisome proliferator-activated receptor)
CC      response element (AB158055) and a minimal promoter; and/or a component
CC      (A2) comprising a nucleic acid encoding a PPAR under control of a
CC      transcriptional promoter. (A), and vectors containing (A1) and (A2), are
CC      used to express (1) in cells for expression of transgenic (1) for
CC      experimental, clinical, therapeutic or diagnostic purposes. (1) encodes
CC      an agriculturally useful, therapeutic, vaccinating or marker protein and
CC      is most especially expressed in human muscle cells. Cells containing (A),
CC      or the vectors, are used to identify PPAR ligands or to produce
CC      transgenic animals for preclinical studies; analysis of bioavailability,
CC      labelling etc. The present sequence is the E/F domain of human
CC      PPARgamma2gamma2, which was used in an example from the invention.
XX
SQ      Sequence 467 AA:

```

```

Query Match          67 0%; Score 2351; DB 22; Length 467;
Best Local Similarity 100.0%; Pred. No. 5.9e-177;
Matches 467; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      222 MNGEKIKFKHITPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQV 281
      1 MNGEKIKFKHITPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVNLIDNDQV 60
Db
QY      282 TLTKYGVHEIITMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFN 341
      61 TLTKYGVHEIITMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEFAVKFN 120
Db
QY      342 ALELDDSLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELOKLNHPRESSOLFATL 401
      121 ALELDDSLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELOKLNHPRESSOLFATL 180
Db
QY      402 LQKMTDLRQIVTEHVOLLQVIAKKTETDMSLHPLQEIYKDYAMAILTGKTTDKSPFIY 461
      181 LQKMTDLRQIVTEHVOLLQVIAKKTETDMSLHPLQEIYKDYAMAILTGKTTDKSPFIY 240
Db
QY      462 DMNSLMGEDKIKFKHITPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVND 521
      241 DMNSLMGEDKIKFKHITPLOESKEVAIRIFQCGQFRSVEAVQETITEYAKSIPGFVND 300
Db
QY      522 LNDQVTLTKYGVHEIITMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEF 581
      301 LNDQVTLTKYGVHEIITMLASLMMKDGVLISEGQGFMTREFLSLRKPGDFMEPKFEF 360
QY      582 AVKFNALELDDSLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELOKLNHPRESSQ 641
      361 AVKFNALELDDSLAIFIAVIIISGDRPGLNVKPIEDIONLLOALELOKLNHPRESSQ 420
Db
QY      642 LFAKLQKMTDLRQIVTEHVOLLQVIAKKTETDMSLHPLQEIYKDY 688
      421 LFAKLQKMTDLRQIVTEHVOLLQVIAKKTETDMSLHPLQEIYKDY 467
Db

```

## RESULT 4

```

AAR93326
ID      AAR93326 standard; Protein: 505 AA.
XX

```

```

AC AAR99326;
XX
XX 14-APR-1997 (first entry)
XX
XX Peroxisome proliferator activated receptor gamma2 subtype 1.
DE
XX Peroxisome proliferator activated receptor gamma; hppargamma; human;
XX peroxisome; cholesterol metabolism; lipid metabolism; respiration;
XX fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
XX herbicide; hppargamma2; adipose tissue; obesity; diabetes; anorexia;
XX hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
XX metabolic disorder; therapy.
XX
XX Homo sapiens.
XX
XX MO9623884-A2.
XX
XX 08-AUG-1996.
XX
XX 29-JAN-1996; 96MO-US01469.
XX
XX 23-OCT-1995; 95US-0005809.
XX 30-JAN-1995; 95US-0380051.
XX 07-JUN-1995; 95US-0484487.
XX
XX (LIGA-) LIGAND PHARM INC.
XX
XX Mukherjee R;
XX
XX WPI; 1996-384114/38.
XX N-PSDB; AAT35334.
XX
XX Human peroxisome proliferator activated receptor(s), hppar-gamma and
XX hppar-gamma2 - useful to identify polypeptide (ant)agonists to treat
XX e.g. obesity, anorexia and diabetes
XX
XX Claim 43; Page 89-92; 108pp; English.
XX
XX AAR99326-R99328 represent the three different subtypes of human
XX peroxisome proliferator activated receptor gamma2 (hppargamma2). These
XX sequences differ in their N-terminal sequences. Peroxisomes contain
XX enzymes for cholesterol and lipid metabolism, and respiration.
XX Peroxisome proliferators increase the capacity of the peroxisomes to
XX metabolise fatty acids, via increased expression of the beta-oxidation
XX cycle enzymes. Peroxisome proliferators include unsaturated fatty
XX acids, hypolipidemic drugs, and herbicides. PPARgamma has two main
XX subtypes (hppargamma, and hppargamma2), which differ in their N-terminal
XX sequences, and are expressed at high levels in adipose tissue.
XX hppargamma proteins repress hpparalpha activity. These proteins can be
XX used in a cell system with a reporter gene (and optionally a hpparalpha
XX protein) to identify agonists and antagonists of the polypeptides, which
XX can be used as therapeutic agents. The agonists and antagonists can be
XX used to treat obesity, diabetes, anorexia, hyperlipidaemia,
XX hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic
XX disorders. These sequences, the DNA encoding them, and antibodies
XX against them, can be used to establish a tissue specific expression
XX pattern for hppargamma.
XX
XX SQ Sequence 505 AA:

```

```

Query Match 63.6%; Score 2231; DB 17; Length 505;
Best Local Similarity 87.7%; Pred. No. 1.9e-167;
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

```

QY 1 MCGTTLDDSPIDPSDFTDLSANISOEMTMVTEMPFPTNGISSVDSLVEDHSHT 60
  |||||||
DB 1 MCGTTLDDSPIDPSDFTDLSANISOEMTMVTEMPFPTNGISSVDSLVEDHSHT 60
QY 61 DIRFTTVSSSTSTPHYEDIPTRNDPVADKYDLKLOEYOSAKVEPASPYSSEKT 120
  |||||||
DB 61 DIRFTTVSSSTSTPHYEDIPTRNDPVADKYDLKLOEYOSAKVEPASPYSSEKT 120
QY 121 QLYN----- 124

```

```

DB 121 QLYNKPHEEPSNSLMAIECVCGDKASGFHYGVHACEGCKGFFRRRTLRKLIYDRCLNC 180
  |||||
QY 125 -----RNCQYCRFOKCLAVGSMHNAIRGRMPQAEKEKLAIESDDIDQLNPESADLR 178
  |||||||
DB 181 RIHKSRNKCQYCRFOKCLAVGSMHNAIRGRMPQAEKEKLAIESDDIDQLNPESADLR 240
QY 179 ALAKHLYDSYIKSFPLTKAKARAILTGKTDKSPFIYOMNSLMGDEDKIKFKHITPLOE 238
  |||||||
DB 241 ALKHLKDYSTIKSFPLTKAKARAILTGKTDKSPFIYIDMNSLMGDEDKIKFKHITPLOE 300
QY 239 QSKKEVARIIFOGCGFSVEAVQETFEYAKSIPGFVLDLNDQYTLKYGHEIITYMLAS 298
  |||||||
DB 301 QSKKEVARIIFOGCGFSVEAVQETFEYAKSIPGFVLDLNDQYTLKYGHEIITYMLAS 360
QY 299 LMKKGVLSSEGCGEMTREFLKLKRPFGDMEPKKEFAVKNALELDDSDLAIFAVI 358
  |||||||
DB 361 LMKKGVLSSEGCGEMTREFLKLKRPFGDMEPKKEFAVKNALELDDSDLAIFAVI 420
QY 359 LSGDRPGLNVKPTIEDIDNLQALELQKLNHPSSQLFAKLLQKMTDLROIVTEHVOL 418
  |||||||
DB 421 LSGDRPGLNVKPTIEDIDNLQALELQKLNHPSSQLFAKLLQKMTDLROIVTEHVOL 480
QY 419 LQYIKKTEYDMSLHPLLOEITKNDLY 443
  |||||||
DB 481 LQYIKKTEYDMSLHPLLOEITKNDLY 505

```

```

RESULT 5
AAEI2868
ID AAEI2868 standard; Protein; 505 AA.
XX
AC AAEI2868;
XX
XX 15-JAN-2002 (first entry)
XX
XX Human peroxisome proliferator activated receptor (PPAR) gamma2 protein.
XX
XX Human; peroxisome proliferator activated receptor gamma2; PPAR gamma2;
XX cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;
XX viral infection; pharmaceutical; thiazolidinedione.
XX
XX Homo sapiens.
XX
XX US6294559-B1.
XX
XX 25-SEP-2001.
XX
XX 03-AUG-1998; 98US-0128142.
XX
XX 02-MAY-1996; 96US-016694P.
XX 18-APR-1997; 97US-0844007.
XX
XX (MERI ) MERCK & CO INC.
XX
XX Smith RG;
XX
XX WPI; 2001-647265/74.
XX N-PSDB; AAD21022.
XX
XX Use of thiazolidinedione for treating cancer and viral infections -
XX
XX Example 3; Fig 4; 17pp; English.
XX
XX The invention relates to compounds and ligands that bind to human
XX peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
XX The invention is useful for treating cancer and other disorders including
XX excessive cell proliferation and viral infection. The invention is also
XX directed to the use of PPAR gamma1 and gamma2 to identify compounds that
XX are antiproliferative, antiviral and antitumour agents. The invention
XX also relates to a method of treating cancer using a pharmaceutical
XX composition comprising thiazolidinedione in an amount sufficient to
XX modulate PPAR gamma1 and gamma2 activity. The present sequence is human

```

CC peroxisome proliferator activated receptor (PPAR) gamma2 protein related  
CC to the invention.

Sequence 505 AA:

Query Match 63.6%; Score 2231; DB 22; Length 505;  
Best Local Similarity 87.7%; Pred. No. 1.9e-167;  
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

QY 1 MGETLGDSPIDPESDSFDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEDHSF 60
   |||
Db 1 MGETLGDSPIDPESDSFDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEDHSF 60
QY 61 DIKPTTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
   |||
Db 61 DIKPTTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
   |||
Db 121 QLYNRPHEEPSNSLMAIECRVCGDKASGFHYGVHACGCGKGFRTIRLKLIDRCDLNC 180
QY 125 -----RNKCYCRFOKCLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 178
   |||
Db 181 RIHKSRMKCOYCRFOKCLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMMGEDKIKFKHTTLPQE 238
   |||
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMMGEDKIKFKHTTLPQE 300
QY 239 OSKEVAIRIFOGCOFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMAS 298
   |||
Db 301 OSKEVAIRIFOGCOFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMAS 360
QY 299 LMNKDGVLSIEGOGFMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSDLAIFIAVII 358
   |||
Db 361 LMNKDGVLSIEGOGFMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSDLAIFIAVII 420
QY 359 LSGDRPGLLVKPIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 418
   |||
Db 421 LSGDRPGLLVKPIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LOVIKKTETDMSLHPLLOEIIYKDL 443
   |||
Db 481 LOVIKKTETDMSLHPLLOEIIYKDL 505

RESULT 6
AAB85800
ID AAB85800 standard; Protein: 505 AA.
XX
AC AAB85800;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PPARgamma protein sequence.
XX
KW PAX8-PPARGamma1; oncogene; cytosstatic; PAX8; PPARgamma1; cancer;
   follicular carcinoma; PPARgamma; human.
XX
OS Homo sapiens.
XX
PN WO200152789-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01664.
XX
PR 20-JAN-2000; 2000US-0177109.
XX
PR 14-AUG-2000; 2000US-0225079.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Krcil TG, Fletcher JA.

```

XX WPI: 2001-514487/56.  
DR N-PSDB; AAH76288.

PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and  
PT treating certain tumors or cancers, e.g. follicular carcinoma  
PS Disclosure: Page 123-125; 145pp; English.

CC The invention relates to an oncogene designated PAX8-PPARGamma1 that  
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The  
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant  
CC methodology. A PPARgamma1 ligand or agent is useful for treating a  
CC subject having a disorder characterized by the presence of a PAX8-  
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.  
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide  
CC and amino acid sequences useful for detecting the above disease. The  
CC present sequence represents a human PPARgamma polypeptide.

Sequence 505 AA:

Query Match 63.6%; Score 2231; DB 22; Length 505;  
Best Local Similarity 87.7%; Pred. No. 1.9e-167;  
Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

```

QY 1 MGETLGDSPIDPESDSFDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEDHSF 60
   |||
Db 1 MGETLGDSPIDPESDSFDTLSANISOEMTWDTMPMPPTNFGISSVDLSVMEDHSF 60
QY 61 DIKPTTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
   |||
Db 61 DIKPTTVDFSSISTPHYEDIPFRTDPVADYKYDKLQEQSAIKVEPASPYYSEKT 120
QY 121 QLYN----- 124
   |||
Db 121 QLYNRPHEEPSNSLMAIECRVCGDKASGFHYGVHACGCGKGFRTIRLKLIDRCDLNC 180
QY 125 -----RNKCYCRFOKCLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 178
   |||
Db 181 RIHKSRMKCOYCRFOKCLAVGSMHNATRFGRMPOAEKEKLLAETSSDIDOLNPESADLR 240
QY 179 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMMGEDKIKFKHTTLPQE 238
   |||
Db 241 ALAKHLVDSYIKSFPLTKAKARAILTGKTTDKSPFVIYDMNSLMMGEDKIKFKHTTLPQE 300
QY 239 OSKEVAIRIFOGCOFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMAS 298
   |||
Db 301 OSKEVAIRIFOGCOFRSYEAVOEITEYAKSIPGFVNLDNDQVTLKYGVEHIIYTMAS 360
QY 299 LMNKDGVLSIEGOGFMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSDLAIFIAVII 358
   |||
Db 361 LMNKDGVLSIEGOGFMTEPEFLKSLRPGDFMEPKFEFAVKNFNALEDDSDLAIFIAVII 420
QY 359 LSGDRPGLLVKPIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 418
   |||
Db 421 LSGDRPGLLVKPIEDIDNLQALELQKLHNPESSQLFAKLLQKMTDLRQIYTEHVQL 480
QY 419 LOVIKKTETDMSLHPLLOEIIYKDL 443
   |||
Db 481 LOVIKKTETDMSLHPLLOEIIYKDL 505

RESULT 7
AAB59839
ID AAB59839 standard; Protein: 505 AA.
XX
AC AAB59839;
XX
DT 27-MAR-2001 (first entry)
XX
DE Human PPARgamma.
XX
KW Cytostatic; antiinflammatory; antisense oligonucleotide; PPARgamma;

```

KW peroxisome proliferator-activated receptor gamma; transcription factor;  
 KW nuclear hormone receptor; human; infection; inflammation; tumour.  
 XX  
 OS Homo sapiens.  
 XX US6159734-A.  
 XX  
 PD 12-DEC-2000.  
 XX  
 PF 18-JAN-2000; 2000US-0484345.  
 XX  
 PR 18-JAN-2000; 2000US-0484345.  
 XX  
 PA (ISIS-) ISIS PHARM INC.  
 XX  
 PI McKay R, Baker BF, Borchers AH;  
 XX  
 DR WPI: 2001-070112/08.  
 DR N-PSDB: AAF23644.  
 XX  
 PT Novel antisense compounds capable of modulating expression of  
 PT peroxisome proliferator-activated receptor gamma useful for diagnosis,  
 PT prophylaxis and treatment of diseases associated with expression of the  
 PT receptor -  
 XX  
 PS Example 13; Column 45-48; 40pp; English.  
 XX  
 CC Peroxisome proliferator-activated receptors (PPARs) are members of the  
 CC nuclear hormone receptor subfamily of transcription factors. The present  
 CC invention relates to antisense oligonucleotides (see AAF23652-F23731),  
 CC targeted to a nucleic acid molecule encoding human PPARgamma, which  
 CC specifically hybridises with and inhibits the expression of human  
 CC PPARgamma. The present sequence is human PPARgamma. The PPARgamma  
 CC antisense oligonucleotides of the present invention can be used in the  
 CC diagnosis and treatment of diseases associated with the expression of  
 CC PPARgamma, e.g. to prevent or delay infection, inflammation or tumour  
 CC formation.  
 CC  
 SQ Sequence 505 AA;  
 Query Match 63.6%; Score 2231; DB 22; Length 505;  
 Best Local Similarity 87.7%; Pred. No. 1.9e-167;  
 Matches 443; Conservative 0; Mismatches 0; Indels 62; Gaps 1;  
 QY 1 MGETLGDSPIDESDSFDTLSANISQEMTWDTMEPMTNFGISSVDLSVMEHSHSF 60  
 DB 1 MGETLGDSPIDESDSFDTLSANISQEMTWDTMEPMTNFGISSVDLSVMEHSHSF 60  
 QY 61 DIKPTTVDFESSISTPHYEDIPFTRTDPVADYKYDKLQEYQSAIKVEPASPYYSEKT 120  
 DB 61 DIKPTTVDFESSISTPHYEDIPFTRTDPVADYKYDKLQEYQSAIKVEPASPYYSEKT 120  
 QY 121 QLYN----- 124  
 DB 121 QLYNKPHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFRRRTIRKLIIYDRCDLNC 180  
 QY 125 -----RNKCOYCRFOCKLAVGSMHNAIRFGMRPOAEKEKLLAETSSDIDOLNPESADLR 178  
 DB 181 RTHKSRNKCQYCRFOCKLAVGSMHNAIRFGMRPOAEKEKLLAETSSDIDOLNPESADLR 240  
 QY 179 ALAKHLDSYIKSPFLTKAKARAILGKTGTDKSPVIYDMNSLMGEGDKIKFKHITPLQE 238  
 DB 241 ALAKHLDSYIKSPFLTKAKARAILGKTGTDKSPVIYDMNSLMGEGDKIKFKHITPLQE 300  
 QY 239 QSKVAIRIRFOCCQFRSVAVOEITETAKSTGFGVNLNDINDOYTLKYGVEHIIITTMLAS 298  
 DB 301 QSKVAIRIRFOCCQFRSVAVOEITETAKSTGFGVNLNDINDOYTLKYGVEHIIITTMLAS 360  
 QY 299 LMKNGVILISEGQGFMTREFLSLRKPPGDFMEPKFEFVKNALFELDDSDLAIFAVI 358  
 DB 361 LMKNGVILISEGQGFMTREFLSLRKPPGDFMEPKFEFVKNALFELDDSDLAIFAVI 420  
 QY 359 LSGDRPGLINVKPIEDIODNLQALELQIKLHPRESSOLFATLQKMTDROIIVTEHVL 418

DB 421 LSGDRPGLINVKPIEDIODNLQALELQIKLHPRESSOLFATLQKMTDROIIVTEHVL 480  
 QY 419 LOVIKKTETDMSLHPLQEIYKLDY 443  
 DB 481 LOVIKKTETDMSLHPLQEIYKLDY 505  
 RESULT 8  
 AAB85801  
 ID AAB85801 standard; Protein; 811 AA.  
 XX  
 AC AAB85801;  
 XX  
 DT 29-OCT-2001 (first entry)  
 XX  
 DE Human PAX8e9(-exon 8)-PPARgammae1 protein sequence.  
 XX  
 KW PAX8-PPARgammae1; oncogene; cytostatic; PAX8; PPARgammae1; cancer;  
 KW follicular carcinoma; PAX8e9(-exon 8)-PPARgammae1; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200152789-A2.  
 XX  
 PD 26-JUL-2001.  
 XX  
 PF 18-JAN-2001; 2001WO-US01664.  
 XX  
 PR 20-JAN-2000; 2000US-0177109.  
 PR 14-AUG-2000; 2000US-0225079.  
 XX  
 PA (BGM) BRIGHAM & WOMENS HOSPITAL INC.  
 PI Kroll TG, Fletcher JA;  
 XX  
 DR WPI: 2001-514487/56.  
 DR N-PSDB: AAH76294.  
 XX  
 PT New PAX8-PPARgammae1 oncogene and oncoprotein, useful for detecting and  
 PT treating certain tumors or cancers, e.g. follicular carcinoma -  
 XX  
 PS Claim 16; Page 127-131; 145pp; English.  
 XX  
 CC The invention relates to an oncogene designated PAX8-PPARgammae1 that  
 CC contains a PAX8 coding region fused to PPARgammae1 coding region. The  
 CC PAX8-PPARgammae1 polypeptides can be expressed by standard recombinant  
 CC methodology. A PPARgammae1 ligand or agent is useful for treating a  
 CC subject having a disorder characterized by the presence of a PAX8-  
 CC PPARgammae1, where the disorder is cancer, e.g. follicular carcinoma.  
 CC The PAX8-PPARgammae1 molecules are also useful for providing nucleotide  
 CC and amino acid sequences useful for detecting the above disease. The  
 CC present sequence represents a human PAX8e9(-exon 8)-PPARgammae1  
 CC polypeptide.  
 CC  
 SQ Sequence 811 AA;  
 Query Match 60.0%; Score 2106.5; DB 22; Length 811;  
 Best Local Similarity 83.7%; Pred. No. 2.4e-157;  
 Matches 425; Conservative 4; Mismatches 14; Indels 65; Gaps 3;  
 QY 1 MGETL-GDSPIDPES--DSFTDTLSANISQEMTWDTMEPMTNFGISSVDLSVMEHSHSF 57  
 DB 304 VGPPLRGYPHPHPTGSGQGSYASSALAGVAEMTWDTMEPMTNFGISSVDLSVMEHSHSF 363  
 QY 58 HSFDIKPTTVDFESSISTPHYEDIPFTRTDPVADYKYDKLQEYQSAIKVEPASPYYSS 117  
 DB 364 HSFDIKPTTVDFESSISTPHYEDIPFTRTDPVADYKYDKLQEYQSAIKVEPASPYYSS 423  
 QY 118 EKTOLYXN----- 124  
 DB 424 EKTOLYXNKPHEEPSNSLMAIECRVCGDKASGFHYGVHACEGCKGFRRRTIRKLIIYDRCD 483

OY	125	-----RRKCOQCFRCQCIANGVSHNRIIRGRNRPQAEKEKLLAEISSDIDLQNPESA	175
DB	484	LNCRTHKSRNKCQYCFRQCLAVGSHNRIIRGRNRPQAEKEKLLAEISSDIDLQNPESA	543
OY	176	DLRALAKHLVDSYIKSPFLTKAKARALITGKTIDKSPFVYTDNNSLMGEDIKFKHTIP	235
DB	544	DLRALAKHLVDSYIKSPFLTKAKARALITGKTIDKSPFVYTDNNSLMGEDIKFKHTIP	603
OY	236	LOEGSRVLAIRIFQCGCFRSVEAVQELTEYAKSIPGFVNLNDQVTLTKGVHEIITYM	295
DB	604	LOEGSRVLAIRIFQCGCFRSVEAVQELTEYAKSIPGFVNLNDQVTLTKGVHEIITYM	663
OY	296	LASLNNKGVLSISGCGFMTREFLKSIRKPGFPMERKFEFAVKFNALELDSDLAIFIA	355
DB	664	LASLNNKGVLSISGCGFMTREFLKSIRKPGFPMERKFEFAVKFNALELDSDLAIFIA	723
OY	356	VIIISGDRPGLLNKPIEDIDNLLQALEIOLKLNHPSSQLFAKLLOKMTDLROIYTEH	415
DB	724	VIIISGDRPGLLNKPIEDIDNLLQALEIOLKLNHPSSQLFAKLLOKMTDLROIYTEH	783
OY	416	VOLLQVYIKTTETDMSLHPLLOEITYKDL	443
DB	784	VOLLQVYIKTTETDMSLHPLLOEITYKDL	811
RESULT 9			
AAB85795			
ID	AAB85795 standard; Protein; 874 AA.		
AC	AAB85795;		
DT	29-OCT-2001 (first entry)		
XX	Human PAX89-PPARGammael protein sequence.		
DE	PAX8-PPARGamma1; oncogene; cytosstatic; PAX8; PPARGamma1; cancer;		
XX	follicular carcinoma; PAX89-PPARGammael; human.		
KW	Homo sapiens.		
OS	wo200152789-A2.		
XX	26-JUL-2001.		
PD	18-JAN-2001; 2001WO-US01664.		
PF	20-JAN-2000; 2000US-0177109.		
PR	14-AUG-2000; 2000US-0225079.		
XX	(BGHM ) BRIGHAM & WOMENS HOSPITAL INC.		
PA	Kro1l TG, Fletcher JA;		
XX	WPI: 2001-514487/56.		
PL	N-PSDB; AAH76283.		
XX	New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and		
PT	treating certain tumors or cancers, e.g. follicular carcinoma		
XX	Claim 16; Page 112-116; 145pp; English.		
PS	The invention relates to an oncogene designated PAX8-PPARGamma1 that		
CC	contains a PAX8 coding region fused to PPARGamma1 coding region. The		
CC	PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant		
CC	methodology. A PPARGamma1 ligand or agent is useful for treating a		
CC	subject having a disorder characterized by the presence of a PAX8-		
CC	PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.		
CC	The PAX8-PPARGamma1 molecules are also useful for providing nucleotide		
CC	and amino acid sequences useful for detecting the above disease. The		
CC	present sequence represents a human PAX89-PPARGammael polypeptide.		
XX	Sequence 874 AA:		

[illegible]

PT treating certain tumors or cancers, e.g. follicular carcinoma -  
XX  
PS Claim 16; page 100-104; 145pp; English.  
XX  
CC The invention relates to an oncogene designated PAX8-PPARgamma1 that  
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The  
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant  
CC methodology. A PPARgamma1 ligand or agent is useful for treating a  
CC subject having a disorder characterized by the presence of a PAX8-  
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.  
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide  
CC and amino acid sequences useful for detecting the above disease. The  
CC present sequence represents a human PAX8-PPARgamma1 polypeptide.  
XX  
SQ Sequence 777 AA:  
  
Query Match 59.8%; Score 2099; DB 22; Length 777;  
Best Local Similarity 83.3%; Pred. No. 8.8e-157;  
Matches 420; Conservative 5; Mismatches 13; Indels 66; Gaps 2;  
  
QY 6 GDSPIPESDSFTDTLSAN----ISOEMTVDTMPMPPTNFGISSVDLSVMEDHSHSPD 61  
DB 274 GKATLPSPNPLGRNLSTHOTYPVVAEMTVDMPTNPFSGISSVDLSVMEDHSHSPD 333  
QY 62 IKPFTVDESSISTPHYEDIPFTPTDPVADYKYDKLQYQSAIKVEPASPYYSEKQ 121  
DB 334 IKPFTVDESSISTPHYEDIPFTPTDPVADYKYDKLQYQSAIKVEPASPYYSEKQ 393  
QY 122 LYN----- 124  
DB 394 LYNKPEEPSLSMAIECRVCGDKSGFHYGHACGCKGFFRTIRLKLIDRCDLNCR 453  
QY 125 -----RNKCOYCRFOKCLAVGSMHNARFGMPQAEKEKLAETSSIDDLNPESADLRA 179  
DB 454 IHKSNRKCQYCRFOKCLAVGSMHNARFGMPQAEKEKLAETSSIDDLNPESADLRA 513  
QY 180 LAHKLDSYTKSPPLTKAKARAILTGKTDKSPVYIYDNLSLMGDEKIKFKHITPLQEQ 239  
DB 514 LAHKLDSYTKSPPLTKAKARAILTGKTDKSPVYIYDNLSLMGDEKIKFKHITPLQEQ 573  
QY 240 SKFAIRIFGCGFRSVEAEOITEYAKSIKPGVNLIDNVOVTLKRGVHEIITYMLASL 299  
DB 574 SKFAIRIFGCGFRSVEAEOITEYAKSIKPGVNLIDNVOVTLKRGVHEIITYMLASL 633  
QY 300 MNKGVILISGCGFMTREFLSLRKPGDEMBKFEFAVKNALDELDDSLAIFIAVIL 359  
DB 634 MNKGVILISGCGFMTREFLSLRKPGDEMBKFEFAVKNALDELDDSLAIFIAVIL 693  
QY 360 SGDRPGILANKPIEDTODNLQALELQKLNHPSSQLPAKLQKMTDLRQIYTEHVOLL 419  
DB 694 SGDRPGILANKPIEDTODNLQALELQKLNHPSSQLPAKLQKMTDLRQIYTEHVOLL 753  
QY 420 QVIKKTETDMSLHPLQETIKDLY 443  
DB 754 QVIKKTETDMSLHPLQETIKDLY 777  
  
RESULT 11  
AAR9323  
ID AAR9323 standard; protein; 494 AA.  
AC AAR9323;  
XX  
DT 14-APR-1997 (first entry)  
XX  
DE Peroxisome proliferator activated receptor gamma subtype 1.  
XX  
KW Peroxisome proliferator activated receptor gamma; PPARgamma; human;  
KW peroxisome; cholesterol metabolism; lipid metabolism; respiration;  
KW fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;  
KW herbicide; hppargamma2; adipose tissue; obesity; diabetes; anorexia;  
KW hyperlipidemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;  
KW metabolic disorder; therapy.

XX  
OS Homo sapiens.  
XX  
PN MO9623884-A2.  
XX  
PD 08-AUG-1996.  
XX  
PF 29-JAN-1996; 96WO-US01469.  
XX  
PR 23-OCT-1995; 95US-0005809.  
PR 30-JAN-1995; 95US-0380051.  
PR 07-JUN-1995; 95US-0484487.  
XX  
PA (LIGA-) LIGAND PHARM INC.  
XX  
PI Mukherjee R;  
XX  
DR WPI; 1996-384114/38.  
DR N-PSDB; AAT35333.  
XX  
PT Human peroxisome proliferator activated receptor(s), hppar-gamma and  
PT hppar-gamma2 - useful to identify polypeptide (antagonists to treat  
e.g. obesity, anorexia and diabetes  
  
Claim 24; page 85-87; 108pp; English.  
  
AAR9323-R99325 represent the three different subtypes of human  
peroxisome proliferator activated receptor gamma (hPPARgamma). These  
sequences differ in their N-terminal sequences. Peroxisomes contain  
enzymes for cholesterol and lipid metabolism, and respiration.  
CC Peroxisome proliferators increase the capacity of the peroxisomes to  
CC metabolize fatty acids, via increased expression of the beta-oxidation  
CC cycle enzymes. Peroxisome proliferators include unsaturated fatty  
CC acids, hypolipidemic drugs, and herbicides. PPARgamma has two main  
CC subtypes (hPPARgamma, and hPPARgamma2), which differ in their N-terminal  
CC sequences, and are expressed at high levels in adipose tissue.  
CC hPPARgamma proteins repress hpparalpha activity. These proteins can be  
CC used in a cell system with a reporter gene (and optionally a hpparalpha  
CC protein) to identify agonists and antagonists of the polypeptides, which  
CC can be used as therapeutic agents. The agonists and antagonists can be  
CC used to treat obesity, diabetes, anorexia, hyperlipidemia,  
CC hypercholesterolaemia, hyperlipoproteinaemia, and other metabolic  
CC disorders. These sequences, the DNA encoding them, and antibodies  
CC against them, can be used to establish a tissue specific expression  
CC pattern for hppargamma.  
XX  
SQ Sequence 494 AA:  
  
Query Match 59.7%; Score 2094; DB 17; Length 494;  
Best Local Similarity 87.0%; Pred. No. 1.2e-156;  
Matches 416; Conservative 0; Mismatches 0; Indels 62; Gaps 1;  
  
QY 28 EMTMVDTMPMPPTNFGISSVDLSVMEDHSHSPDIPFTTVPDSSISTPHYEDIPFTRID 87  
DB 17 EMTMVDTMPMPPTNFGISSVDLSVMEDHSHSPDIPFTTVPDSSISTPHYEDIPFTRID 76  
QY 88 PNVADYKYDKLQYQSAIKVEPASPYYSEKQLYN----- 124  
DB 77 PNVADYKYDKLQYQSAIKVEPASPYYSEKQLYN----- 124  
QY 125 -----RNKCOYCRFOKCLAVGSMHNA 145  
DB 137 GFHYGVHACGCKGFFRTIRLKLIDRCDLNCRHKKSRNKQYCRFOKCLAVGSMHNA 196  
QY 146 IRGRMPQAEKEKLAETSSIDDLNPESADLRALKHLYDSYTKSPPLTKAKARAILTG 205  
DB 197 IRGRMPQAEKEKLAETSSIDDLNPESADLRALKHLYDSYTKSPPLTKAKARAILTG 256  
QY 206 KTDKSPFYIYDNLSLMGDEKIKFKHITPLQEQSKVAIRIFGCGFRSVEAEOITEY 265  
DB 257 KTDKSPFYIYDNLSLMGDEKIKFKHITPLQEQSKVAIRIFGCGFRSVEAEOITEY 316  
QY 266 AMSIPGVNLIDNVOVTLKRGVHEIITYMLASLMNKDGLISEGCGFMTREFLSLRK 325

```

Db 317 AKSIPGFVNLDNDQVTLTKYGVHEIITYMLASLNKNGDVLISEGQGMTRFEFLSKRP 376
Oy 326 FGDFMEPKFEFAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPIRIDIDONLQALEL 385
Db 377 FGDFMEPKFEFAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPIRIDIDONLQALEL 436
Oy 386 QLKLNHPRESSQLFAKLQKMTDLROIVTEHVQLQVIRKKTETDMSLHPLQETIKDLY 443
Db 437 QLKLNHPRESSQLFAKLQKMTDLROIVTEHVQLQVIRKKTETDMSLHPLQETIKDLY 494

RESULT 12
AAB85794
ID AAB85794 standard; Protein: 840 AA.
XX
AC AAB85794;
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PAX8e8-PPARGammae1 protein sequence.
XX
KW PAX8-PPARGammae1; oncogene; cytosolic; PAX8; PPARGammae1; cancer;
KW follicular carcinoma; PAX8e8-PPARGammae1; human.
XX
OS Homo sapiens.
XX
PN WO200152789-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01664.
XX
PR 20-JAN-2000; 2000US-0177109.
XX
PR 14-AUG-2000; 2000US-0225079.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Kroll TG, Fletcher JA:
XX
PI WPI: 2001-514487/56.
XX
DR N-PSDB; AAB76282.
XX
PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
XX
PS Claim 16; Page 106-109; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARGammae1 that
CC contains a PAX8 coding region fused to PPARGammae1 coding region. The
CC PAX8-PPARGammae1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGammae1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGammae1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGammae1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e8-PPARGammae1 polypeptide.
XX
XX
Sequence 840 AA:
Oy Query Match 59.7%; Score 2094; DB 22; Length 840;
Best Local Similarity 87.0%; Pred. No. 2,4e-156;
Matches 416; Conservative 0; Mismatches 0; Indels 62; Gaps 1:

```

```

Db 483 GFHYGVHACEGCKGFRRRTIRKLTIYDRCDLNCRIHKKSRNKCQCRQKCLAVGMSHNA 542
Oy 146 IRFGMPQAEKEKLAIESSDIDQNPESADLRALAKLYDSYISFPLTAKARAILTG 205
Db 543 IRFGMPQAEKEKLAIESSDIDQNPESADLRALAKLYDSYISFPLTAKARAILTG 602
Oy 206 KTTDPSPIVIDMNSLMNGEDKIKFKHITPLOEGSKEVAIRFQGCGRSEAVAOETREY 265
Db 603 KTTDPSPIVIDMNSLMNGEDKIKFKHITPLOEGSKEVAIRFQGCGRSEAVAOETREY 662
Oy 266 AKSIPGFVNLDNDQVTLTKYGVHEIITYMLASLNKNGDVLISEGQGMTRFEFLSKRP 325
Db 663 AKSIPGFVNLDNDQVTLTKYGVHEIITYMLASLNKNGDVLISEGQGMTRFEFLSKRP 722
Oy 326 FGDFMEPKFEFAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPIRIDIDONLQALEL 385
Db 723 FGDFMEPKFEFAVKFNALIEDSDLAIFIAVITILSGDRPGLLNKPIRIDIDONLQALEL 782
Oy 386 QLKLNHPRESSQLFAKLQKMTDLROIVTEHVQLQVIRKKTETDMSLHPLQETIKDLY 443
Db 783 QLKLNHPRESSQLFAKLQKMTDLROIVTEHVQLQVIRKKTETDMSLHPLQETIKDLY 840

RESULT 13
AAR99324
ID AAR99324 standard; Protein: 477 AA.
XX
AC AAR99324;
XX
DT 14-APR-1997 (first entry)
XX
DE Peroxisome proliferator activated receptor gamma subtype 2.
XX
KW Peroxisome proliferator activated receptor gamma; PPARGamma; human;
KW peroxisome; cholesterol metabolism; lipid metabolism; respiration;
KW fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
KW hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
KW metabolic disorder; therapy.
XX
OS Homo sapiens.
XX
PN WO9623884-A2.
XX
PD 08-AUG-1996.
XX
PF 29-JAN-1996; 96WO-US01469.
XX
PR 23-OCT-1995; 95US-0005809.
XX
PR 30-JAN-1995; 95US-0380051.
XX
PR 07-JUN-1995; 95US-0484487.
XX
PA (LIGA-) LIGAND PHARM INC.
XX
PI Mukherjee R:
XX
PI WPI: 1996-384114/38.
XX
DR N-PSDB; AAT35333.
XX
PT Human peroxisome proliferator activated receptor(s), hPPAR-gamma and
PT hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat
PT e.g. obesity, anorexia and diabetes
XX
PS Claim 25; : 108pp; English.
XX
XX
AAR99323-R99325 represent the three different subtypes of human
CC peroxisome proliferator activated receptor gamma (hPPARGamma). These
CC sequences differ in their N-terminal sequences. Peroxisomes contain
CC enzymes for cholesterol and lipid metabolism, and respiration.
CC Peroxisome proliferators increase the capacity of the peroxisomes to
CC metabolise fatty acids, via increased expression of the beta-oxidation
CC cycle enzymes. Peroxisome proliferators include unsaturated fatty

```





Db 361 GDFMEPKFEFAVKFNALDELDDSLAIFIAVITLLSGDRGGLNVKPIEDIQDNLQALELQ 420  
QY 387 LKLNHPSSQLFAKLQKMTDLROIVTEHVOLLQYIKKTEMDSLHPLQEIYKDL 443  
DB 421 LKLNHPSSQLFAKLQKMTDLROIVTEHVOLLQYIKKTEMDSLHPLQEIYKDL 477

RESULT 15  
AAE12867  
ID AAE12867 standard; Protein: 477 AA.  
XX  
AC AAE12867;  
XX  
DT 15-JAN-2002 (first entry)  
XX  
DE Human peroxisome proliferator activated receptor (PPAR) gamma1 protein.  
XX  
KW Human; peroxisome proliferator activated receptor gamma1; PPAR gamma1;  
KW cytochrome; antiproliferative; antiviral; cancer; cell proliferation;  
KW viral infection; pharmaceutical; thiazolidinedione.  
XX  
OS Homo sapiens.  
XX  
PM US6294559-B1.  
XX  
PD 25-SEP-2001.  
XX  
PF 03-AUG-1998; 98US-0128142.  
XX  
PR 02-MAY-1996; 96US-016694P.  
PR 18-APR-1997; 97US-0844007.  
XX  
PA (MERI ) MERCK & CO INC.  
XX  
PI Smith RG;  
XX  
DR WPI: 2001-647265/74.  
DR N-PSDB; AAD21021.  
XX  
XX  
PT Use of thiazolidinedione for treating cancer and viral infections -  
XX  
PS Example 3; Fig 2; 17pp; English.  
XX  
CC The invention relates to compounds and ligands that bind to human  
CC peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.  
CC The invention is useful for treating cancer and other disorders including  
CC excessive cell proliferation and viral infection. The invention is also  
CC directed to the use of PPAR gamma1 and gamma2 to identify compounds that  
CC are antiproliferative, antiviral and antitumor agents. The invention  
CC also relates to a method of treating cancer using a pharmaceutical  
CC composition comprising thiazolidinedione in an amount sufficient to  
CC modulate PPAR gamma1 and gamma2 activity. The present sequence is human  
CC peroxisome proliferator activated receptor (PPAR) gamma1 protein related  
CC to the invention.  
XX  
SQ Sequence 477 AA;

Query Match 59.5%; Score 2089; DB 22; Length 477;  
Best Local Similarity 87.0%; Pred. No. 2.8e-156;  
Matches 415; Conservative 0; Mismatches 0; Indels 62; Gaps 1;

QY 29 MTMVDTEPFWPTNGISSVDLSVMEHSHSFDIKPFTTVDFSSISSTPHYEDIPFTTDP 88  
DB 1 MTMVDTEPFWPTNGISSVDLSVMEHSHSFDIKPFTTVDFSSISSTPHYEDIPFTTDP 60  
QY 89 VVADKYDLKLOEYOSAIKVPASPPYSEKTOYLN----- 124  
DB 61 VVADKYDLKLOEYOSAIKVPASPPYSEKTOYLNKPHPEPSNSLMAIECVCGDKASG 120  
QY 125 -----RNKCYCRPOKCLAVGSHNAI 146  
DB 121 FHVGHAGEGCGFFRRITRIKLITDRCDLMCRHKSRNKCQYCRFOKCLAVGSHNAI 180

QY 147 RFGMPQAEKEKLLAEISSDIDQNPESADLRALAKHLVDSYIKSPPLYTKAKARAILTGK 206  
DB 181 RFGMPQAEKEKLLAEISSDIDQNPESADLRALAKHLVDSYIKSPPLYTKAKARAILTGK 240  
QY 207 TTDSPPFIYDMNSLMGKEDKIKRKHITPLOBOSKEVAIRIFOGCQFRSVAVOEITEYA 266  
DB 241 TTDSPPFIYDMNSLMGKEDKIKRKHITPLOBOSKEVAIRIFOGCQFRSVAVOEITEYA 300  
QY 267 KSIPGFVNLNDQVTLKYGVHEIITYMLASLNNKDGVLISEGQGFMTREFLSLRKF 326  
DB 301 KSIPGFVNLNDQVTLKYGVHEIITYMLASLNNKDGVLISEGQGFMTREFLSLRKF 360  
QY 327 GDFMEPKFEFAVKFNALDELDDSLAIFIAVITLLSGDRGGLNVKPIEDIQDNLQALELQ 386  
DB 361 GDFMEPKFEFAVKFNALDELDDSLAIFIAVITLLSGDRGGLNVKPIEDIQDNLQALELQ 420  
QY 387 LKLNHPSSQLFAKLQKMTDLROIVTEHVOLLQYIKKTEMDSLHPLQEIYKDL 443  
DB 421 LKLNHPSSQLFAKLQKMTDLROIVTEHVOLLQYIKKTEMDSLHPLQEIYKDL 477

Search completed: February 25, 2003, 03:58:26  
Job time : 65 secs

**THIS PAGE BLANK (USPT)**

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 25, 2003, 03:59:06 ; Search time 368 Seconds  
(without alignments)  
4210.258 Million cell updates/sec

Title: US-09-931-007A-1

Perfect score: 3508  
Sequence: 1 MGRTLDSPIDPESDSFTD.....KTEEDMSIHLIDYKLDY 688

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
MODEL=frame+ p2n.model -DEV=xlh  
-O=/cgn2\_1/USPRO.spool/US09931007/rnat\_14022003\_102553\_4668/app\_query.fasta\_1.839  
-DB=/LGeneseq\_101002 -QFRT=fastcap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR\_SCORE=pct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=us09931007 @cgn.1.1.236 @rnat.14022003.102553.4668 -NCPU=6 -ICPU=3  
-NO\_XLPHY -NO\_MMAP -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
1: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SID2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SID2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES					
Result No.	Score	Query Match	Length	DB ID	Description
1	2231	63.6	1518	22 AAD21022	Human peroxisome p
2	2231	63.6	1608	22 AAH76288	Human PPARgamma cd
3	2231	63.6	1608	22 AAF3644	Human PPARgamma co
4	2231	63.6	1647	22 AAT35334	Peroxisome prolif
5	2106.5	60.0	2596	22 AAH76294	Human PAX8b(- exon
6	2106.5	60.0	2625	22 AAH76283	Human PAX8b/-PPARg
7	2099	59.8	2334	22 AAH76281	Human PAX8b/-PPARg
8	2098	59.8	1936	22 AAT35333	Peroxisome prolif
9	2094	59.7	2523	22 AAH76282	Human PAX8b/-PPARg
10	2089	59.5	1434	22 AAD21021	Human peroxisome p
11	2065.5	58.9	1811	22 AAH76292	Human PPAR gamma1
12	2065.5	58.9	1811	22 AAH76296	Human PPARgamma cd
13	2065.5	58.9	1811	22 AABN95714	Gene #2212 used to
14	2065.5	58.9	1844	22 AAF55663	Human peroxisome p
15	2065.5	58.9	2295	22 AAS44653	Human full-length
16	2065	58.9	2005	22 AAX35481	Human peroxisome p
17	2065	58.9	2005	17 AAT16205	Peroxisome prolif
18	2065	58.9	2005	20 AAX39814	Mouse peroxisome p
19	1570	44.8	1278	22 AAH33499	Human colon cancer
20	1101.5	31.4	1326	17 AAT10583	Peroxisome prolif
21	1101.5	31.4	3294	14 AAQ48939	Steroid hormone re
22	1101.5	31.4	3301	20 AAX36523	Human PPAR-delta c
23	1101.5	30.8	1850	22 AAF30517	Lung cancer relate
24	1079.5	30.7	1407	16 AAQ92081	Peroxisome prolif
25	1076.5	30.7	2012	17 AAT16206	Human peroxisome p
26	1076.5	30.7	2012	20 AAX39815	Mouse peroxisome p
27	1076.5	30.7	1854	20 AAX36521	Human PPAR-alpha c
28	1075.5	30.7	2009	14 AAQ39089	XR4 coding sequenc
29	1056	30.1	732	20 AAX19070	Human colon cancer
30	761	21.7	498	24 ABO59148	Human colon cancer
31	618.5	17.6	415	19 AAV398246	Human colon cancer
32	562	16.0	373	21 AAV21207	Mouse PPAR-gamma A
33	521	14.9	313	21 AAC98099	Human colon cancer
34	512.5	14.6	611	21 AAC98099	Human colon cancer
35	486	13.9	589	24 ABO60211	Human breast cell
36	472	13.5	451	22 ABA48744	Human foetal liver
37	472	13.5	451	22 ABA66654	Probe #12185 for g
38	472	13.5	451	22 ABA33719	Human brain expres
39	472	13.5	451	22 AAK15083	Human bone marrow
40	472	13.5	451	22 AAK08811	Probe #11514 for g
41	472	13.5	451	22 AAT1581	Probe #15556 used
42	472	13.5	451	22 AAI6870	Probe #7266 used t
43	472	13.5	451	22 AAI07275	Human genome-deriv
44	472	13.5	451	24 ABS14765	Human colon cancer
45	435	12.4	576	24 ABO57991	
ALIGNMENTS					
RESULT 1					
AAD21022					
AAD21022 standard; CDNA: 1518 BP.					
AAD21022;					
15-JAN-2002 (first entry)					
XX	Human peroxisome proliferator activator receptor (PPAR) gamma2 CDNA.				
XX	Human: peroxisome proliferator activator receptor gamma2; PPAR gamma2;				
KW	cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;				
KM	viral infection; pharmaceutical; thiazolidinedione; ss.				
XX					
OS	Homo sapiens.				
XX					
XX	Key	Location/Qualifiers			
FT	CDS	1..1518			

```

FT      /*tag= a
FT      /product= "Human peroxisome proliferator activated
FT      receptor (PPAR) gamma2 protein"
PN      US6294559-B1.
XX      25-SEP-2001.
XX      PD
XX      03-AUG-1998; 98US-0128142.
XX      PF
XX      02-MAY-1996; 96US-016694P.
XX      PR 18-APR-1997; 97US-0844007.
XX      PA (MERI ) MERCK & CO INC.
XX      PI Smith RG;
XX      DR WPI: 2001-647265/74.
XX      DR P-PSDB: AAE12868.
PT      Use of thiazolidinedione for treating cancer and viral infections -
PS      Example 3; Fig 3; 17pp; English.
XX
XX      The invention relates to compounds and ligands that bind to human
CC      peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
CC      The invention is useful for treating cancer and other disorders including
CC      excessive cell proliferation and viral infection. The invention is also
CC      directed to the use of PPAR gamma1 and gamma2 to identify compounds that
CC      are antiproliferative, antiviral and antitumor agents. The invention
CC      also relates to a method of treating cancer using a pharmaceutical
CC      composition comprising thiazolidinedione in an amount sufficient to
CC      modulate PPAR gamma1 and gamma2 activity. The present cDNA sequence
CC      encodes human peroxisome proliferator activated receptor (PPAR) gamma2
CC      protein related to the invention.
XX
XX      Sequence 1518 BP; 438 A; 363 C; 346 G; 371 T; 0 other:

Alignment Scores:
Pred. No.: 1,74e-197 Length: 1518
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: Gaps: 1

US-09-931-007A-1 (1-688) x AAD21022 (1-1518)
OY      1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
DB      1 ATGGGTGAAACTCTGGAGATTCTCTATTGACCCAGAAAGCGATTCTTCACATGATACA 60
OY      21 LeuSerAlaAsnIleSerGlnIleMetThrMetValAspThrGluMetProPheThrPro 40
DB      61 CTGTGTGCAAAATATATCAAGAAATGACCATGTTGACACAGAGATGCGCATTTGGCCC 120
OY      41 ThrAsnPhgGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60
DB      121 ACCAACTTTGGATCAGCTCCGTCGATCTCTCCGTAATGAGAACACCCACCTCCTTT 180
OY      61 AspIleLeuSerProPheThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
DB      181 GATATCAAGCCCTTCACACTGTTGACTTCTCCACATTCTTACATCCACATPACGAAGAC 240
OY      81 IleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeuIleGln 100
DB      241 ATTCATTTCACAGAACACGATCCAGTGTGCGAGATTCAAAATAGACCTGAACCTTCAA 300
OY      101 GluTyrGlnSerAlaIleLeuValGluProAlaSerProTyrTyrSerGluTyrStr 120
DB      301 GAGTACCAAGAGCAATCAATGAGAGCTGCATCTCCACCTTATATTTCAGAAAGACT 360
OY      121 GlnLeuTyrAsn----- 124

```

```

DB      361 CAGCTCTAACAATAAGCCTCATGAGAGCCTTCCAACTCCCTCATGGAATGATGTCT 420
OY      124 ----- 124
DB      421 GTCCTGAGATTAAGCTTCTCGAATTTCTACTATGAGAGTTCAATGCTTGTGAAGATGCAAG 480
OY      124 ----- 124
DB      481 GGTTCCTTCGGAAGAACATCAATGAAAGCTTATCTATGACAGATGTGATCTTAAGTGT 540
OY      125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
DB      541 CGGATCCACAAAAAAGTGAATTAAGTCACTACTGTCGCTTCAGAAATAGCTTCA 600
OY      139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
DB      601 GTGGGATGCTCTCATTAATGCCATACAGGTTTGGCGGATGCCACAGCCGAGAAAGAGAG 660
OY      159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178
DB      661 CTGTTGGCGGAGATCTCCAGTATATCCAGCTGAATCCAGAGTCCGCTACCTCCGG 720
OY      179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198
DB      721 GCCCTGGCAAAACATTTGATGACTCATATAAAGTCTTCCGTCGACCAAAAGCAAG 780
OY      199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218
DB      781 GCGAGGGCATCTTACAGGAAAGCAACAGCAATACCACTTCCTTATCTATGACATG 840
OY      219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnLys 238
DB      841 AATTCCTTAATGATGGAGAAATCAATCAAGTTCAAAACATGACACCCCTCGAGAG 900
OY      239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
DB      901 CAGACCAAGAGAGTGCCATCCGATCTTTCAGGGCTCCCAATTCGCTCCGTGAGACT 960
OY      259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
DB      961 GTCCAGAGATACAGAGATGATGCCAAAAAGCATCTCTGGTTTGTAAATCTTGACTTGAC 1020
OY      279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
DB      1021 GACCAAGTAACCTCTCCCAATATGAGAGCCACGAGATCATTTACACAAATGCTGCCCTCC 1080
OY      299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
DB      1081 TTGATGAATTAAGATGGGGTCTCTCATATCCGAGGCCAAGGCTTCATGACAAAGGAGTTT 1140
OY      319 LeuLysSerLeuArgLysProPheGlyAspPheMetGlnProLysPheGluPheAlaVal 338
DB      1141 CTTAAAGAGCTGCGAAGGCTTTTGTGACTTTATGAGAGCCCAAGTTTGTGAGTTGCTGTG 1200
OY      339 LysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle 358
DB      1201 AAGTTCATATGCACCTGGATTTGATGATGACAGCGACTTGGCAATATTATTTGCTATTATT 1260
OY      359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
DB      1261 CTCAGTGGAGACCGGCCGAGGTTTGTGTAATGGAAGCCCATTTGAAGACATTCAGAACAC 1320
OY      379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPhe 398
DB      1321 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGCTGAAGCCACCTGAGTCTCCACAGCTGTT 1380
OY      399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeu 418
DB      1381 GCCAAGCTGCTCCAGAAATATGACAGACTCAGACAGATTTGACAGGAGACAGTGCAGCTA 1440
OY      419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438

```

Db 1441 CTGACGATGATCAAGACGAGACAGACATGACTTTCACCCGCTCCTGAGAGATC 1500  
Qy 439 TyrlYsAspLeuTyr 443  
Db 1501 TACAGGACTTGTAC 1515  
RESULT 2  
AAH76288  
ID AAH76288 standard; cDNA; 1608 BP.  
XX  
AC AAH76288;  
XX  
DT 29-OCT-2001 (first entry)  
XX  
DE Human PPARgamma cDNA sequence.  
XX  
KM PAX8-PPARgamma1; oncogene; cytosolic; PAX8; PPARgamma1; cancer;  
KM follicular carcinoma; PPARgamma; human; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 91..1608  
FT /tag= a  
XX  
MO200152789-A2.  
XX  
PD 26-JUL-2001.  
XX  
PE 18-JAN-2001; 2001MO-US01664.  
XX  
PR 20-JAN-2000; 2000US-0177109.  
PR 14-AUG-2000; 2000US-0225079.  
XX  
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
XX  
PI Kroll TG, Fletcher JA;  
XX  
DR WPI: 2001-514487/56.  
DR P-PSDB; AAB85800.  
XX  
PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and  
PT treating certain tumors or cancers, e.g. follicular carcinoma -  
XX  
PS Disclosure: Page 123-125; 145pp; English.  
XX  
CC The invention relates to an oncogene designated PAX8-PPARgamma1 that  
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The  
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant  
CC methodology. A PPARgamma1 ligand or agent is useful for treating a  
CC subject having a disorder characterized by the presence of a PAX8-  
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.  
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide  
CC and amino acid sequences useful for detecting the above disease. The  
CC present sequence represents a human PPARgamma1 polypeptide encoding cDNA.  
XX  
SQ Sequence 1608 BP; 462 A; 380 C; 361 G; 405 T; 0 other;  
Alignment Scores:  
Pred. No.: 1,87e-197 Length: 1608  
Score: 2231.00 Matches: 443  
Percent Similarity: 87.72% Conservative: 0  
Best Local Similarity: 87.72% Mismatches: 0  
Query Match: 63.60% Indels: 62  
DB: 22 Gaps: 1  
US-09-931-007a-1 (1-688) x AAH76288 (1-1608)  
Qy 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20  
Db 91 ATGGGTAAACTCTGGGAGATTCTCTATTGACCAGCAAGCGATTCCTTACTGATACCA 150  
Qy 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheTrpPro 40

Db 151 CTGTCGCACAAACATATATCACAAGAAATGACCATGGTTGACACAGAGATGCCATTCCTGGCCC 210  
Qy 41 ThrAsnPhcGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisPhe 60  
Db 211 ACCAACTTGGGATCAGCTCCGCTGATCTCTCCGTAATGGAAGACCACTCCACCTCTTT 270  
Qy 61 AspIleLysProPheThrThrValAspPheSerIleSerThrProHisTyrGluAsp 80  
Db 271 GATATCAAGCCCTTCACTACTACTGTGACTTCTCCAGCATTTTCTACTCCCATTTAGCAAGAC 330  
Qy 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
Db 331 ATTCATTCACAGAACAGATCCAGTGGTTCAGATTACAGTATGACCTGAACCTTCAA 390  
Qy 101 GluTyrGlnSerAlaIleLysValGluProLaseRPropotyrrTyrSerGluLysThr 120  
Db 391 GAGTACCAAGTGCATTCAAAGTGGAGCTGCATCTCCACCTTATTTATTCGAGAAGACT 450  
Qy 121 GlnLeuTyrAsn----- 124  
Db 451 CAGCTCTACAAATAAGCCTCATGAAGAGCCTTCCAACTCCCTCATGGCAATTGAATGTCGT 510  
Qy 124 ----- 124  
Db 511 GTCGTGGAGATAAAGCTTCTGCAATTCATGAGATTCATGCTGTGGAAGATGACAG 570  
Qy 124 ----- 124  
Db 571 GGTTCCTCCGAGAACAAATCAGATTCGAAGCTTATCTATGACAGATGATCTTAACGT 630  
Qy 125 -----ArgAsnLysCysGlnIrrCysArgPheGlnLysCysLeuAla 138  
Db 631 CGATCCACAAAAAAGTGAATTAATGTCAGTCTCGGTTTCACAAATGCCCTTGCA 690  
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158  
Db 691 GTGGGATGCTCATTAATGCCATCAGTTTGGCGGATGCCAGCCGAGCAAGAGAGAG 750  
Qy 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg 178  
Db 751 CTGTTGGCGGAGACTCCAGTGAATATCCAGCTGAATCCAGATCGCGTACCTCCGG 810  
Qy 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198  
Db 811 GCCCTGCCAAATATTTGTTATGACTCATACATAAGTCTTCCGCTGACCAAGACAAAG 870  
Qy 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
Db 871 GCGAGGGGATCTTGACAGGAAGACACAGACAAATACCATTCGTTATCTATGACATG 930  
Qy 219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu 238  
Db 931 AATTCCTTAAAGATGGGAGAGATTAATCAAGTTCAAAACATCACCCCTCGAGAG 990  
Qy 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258  
Db 991 CAGAGCAAAAGAGTGGCCATCCGATCTTTTACGGGCTGCCAGTTTCCGCTCGTAGAGCT 1050  
Qy 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278  
Db 1051 GTGCAGAGGATCACAGAGTATGCCAAAGCATTCCTGTTTGTAAATCTTGACTGATGAC 1110  
Qy 279 AspGlnValThrLeuLeuLysTyrGlyAlaHisGluIleIleTyrThrMetLeuAlaSer 298  
Db 1111 GACCAAGTAACTCTCCCAATATGAGATCCACAGATCATTTACAAATCTCGGCTCC 1170  
Qy 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318  
Db 1171 TTGATGAATTAAGATGGGTTCTCATATCCGAGGGCCAAAGCTTCATGACAGGAGATT 1230  
Qy 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338

Db 1231 CTAAGACCTCGGAAGCCTTTGGTGACTTTATGAGACCCCAAGTTTGATTTGCTGTG 1290  
OY 339 LysPheAsnAlaLeuGlnLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358  
Db 1291 AAGTTCATGCGACTGGAAATGATGACAGCGACTGGCAATTTTATTTGCTCATTTAT 1350  
OY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378  
Db 1351 CTCAGTGGAGACCGCCGAGCTTTGCTGAATGTGAAGCCATTGAAGACATTCAGACAAAC 1410  
OY 379 LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPhe 398  
Db 1411 CTGCTACAGACCCCTGGAGCTCCAGCTGAAGCTGAACCCCTGAGCTCCACAGCTGT 1470  
OY 399 AlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418  
Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTGTACGGAAACAGCTGACGTA 1530  
OY 419 LeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIle 438  
Db 1531 CTGCAAGTGATCAGAGAGCGAGACAGACATGATCTTCACCCGCTCCTGCAGAGATC 1590  
OY 439 TyrLysAspLeuTyr 443  
Db 1591 TACAAGACTTGTAC 1605  
RESULT 3  
AAF23644  
ID AAF23644 standard; cDNA; 1608 BP.  
XX AAF23644;  
AC 27-MAR-2001 (first entry)  
XX  
XX  
XX  
XX  
DE Human PPARgamma coding sequence.  
XX  
XX Cytostatic; antiinflammatory; antisense oligonucleotide; PPARgamma;  
KW peroxisome proliferator-activated receptor gamma; transcription factor;  
KW nuclear hormone receptor; human; infection; inflammation; tumour; ss.  
XX  
XX OS Homo sapiens.  
XX  
XX US6159734-A.  
XX  
XX 12-DEC-2000.  
XX  
XX 18-JAN-2000; 2000US-0484345.  
XX  
XX 18-JAN-2000; 2000US-0484345.  
XX  
XX (ISIS-) ISIS PHARM INC.  
XX  
XX McKay R, Baker BF, Borchers AH;  
XX  
XX WPI: 2001-070112/08.  
XX  
XX P-PSDB: AAB59839.  
XX  
XX Novel antisense compounds capable of modulating expression of  
PT peroxisome proliferator-activated receptor gamma useful for diagnosis,  
PT prophyllaxis and treatment of diseases associated with expression of the  
PT receptor -  
XX  
XX Example 13; Column 45-48; 40pp; English.  
XX  
XX Peroxisome proliferator-activated receptors (PPARs) are members of the  
XX nuclear hormone receptor subfamily of transcription factors. The present  
XX invention relates to antisense oligonucleotides (see AAF23652-F23731),  
XX targeted to a nucleic acid molecule encoding human PPARgamma, which  
XX specifically hybridises with and inhibits the expression of human  
XX PPARgamma. The present sequence is the coding sequence for human  
XX PPARgamma. The PPARgamma antisense oligonucleotides of the present  
XX invention can be used in the diagnosis and treatment of diseases  
XX associated with the expression of PPARgamma, e.g. to prevent or delay

CC infection, inflammation or tumour formation.  
XX  
SQ Sequence 1608 BP; 462 A; 380 C; 361 G; 405 T; 0 other;  
Alignment Scores:  
Pred. No.: 1,87e-197 Length: 1608  
Score: 2231.00 Matches: 443  
Percent Similarity: 87.72% Conservative: 0  
Best Local Similarity: 87.72% Mismatches: 0  
Query Match: 63.60% Indels: 62  
Db: 22 Gaps: 1  
US-09-931-007a-1 (1-688) x AAF23644 (1-1608)  
OY 1 MetGlyGlnThrLeuGlnLysAspSerProIleAspProGlnUserAspSerPheThrAspThr 20  
Db 91 ATGGGTGAACCTCTGGAGATTTCTCTATTGACCCAGAAAGGATTCCTTCACTGATACA 150  
OY 21 LeuSerAlaAsnIleSerGlnLysMetThrMetValAspThrGlnMetProPheThrPro 40  
Db 151 CTGCTGCAAAACATATCAACAAGAAATGACATGTTGACACAGAGATGCATTTGCGCC 210  
OY 41 ThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPhe 60  
Db 211 ACCAAGCTTTGGGATCAGCTCCGCTGATCTCTCCGTAATGGAAGACCACTCCACCTCTT 270  
OY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGlnAsp 80  
Db 271 GATATCAACCCCTTCACTACAGTCTTCACTTCTCCACATTTTCACTCCACATTTAGCAAGAC 330  
OY 81 IleProPheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGln 100  
Db 331 ATTCATTCACAAAGAACAGATCCAGTGCAGATTACAGTATGACGCTGAACCTTCAA 390  
OY 101 GlnTyrGlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGlnLysThr 120  
Db 391 GAGTACCAAAAGTCATCAATCAAAAGTGGAGCCGTGATCTCCACCTTATATTCTCAGAAACT 450  
OY 121 GlnLeuTyrAsn----- 124  
Db 451 CAGCTCTACAAATACCCCTCATGAAGACCTTCCAACTCCCTCATGGCAATTGAATGTGT 510  
OY 124 ----- 124  
Db 511 GTCTGTGAGATAAAGCTTCTGATTTCACTATGAGATTGATGTTGAAGATGCAAG 570  
OY 124 ----- 124  
Db 571 GGTTCCTCCGAGAACATCAATGAGATTGAAGCTTATCTATGACAGATGTGATCTTAACGT 630  
OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
Db 631 CGGATCCACAAAAGAAAGTAATGATGTCAGTCTGTGCTTTCAGAAATGCTTGTCA 690  
OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLys 158  
Db 691 GTGGGATGCTCATTAATGCCATCAGTTTGGGCGGATGCCACAGCGGAGAGAGAAAG 750  
OY 159 LeuLeuAlaGlnIleLysSerSerAspIleAspGlnLeuAsnProGlnUserAlaAspLeuArg 178  
Db 751 GTTTGGCGGAGATCTCCAGTGTATATGACCAAGCTGAATCCAGATCCGCTGACCTCCGG 810  
OY 179 AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys 198  
Db 811 GCCCTGGCAAAACATTTGTATGATCATATCAATGAAGTCTTCCCTGCCCTCAACAAAG 870  
OY 199 AlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMet 218  
Db 871 GCGAGGCGGATTTGACAGGAAGAACACAGACAAATCACCATTGCTATCATATGACATG 930  
OY 219 AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGln 238  
Db 931 AATTCCTTAATGATGGAGAGAAAGATAAATCAAGTTCAAAACATCACCCTCCCTGCAGAG 990

```

QY 239 GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla 258
    |||
Db 991 CAGACCAAAAGAGGTGGCCATCCGCATCTTTCAGGGCTGCGCATTCCTCCCTGGAGGCT 1050
QY 259 ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn 278
    |||
Db 1051 GTCCAGAGATCAGACAGATATGCCAAAAGCATTTCTGGTTTGTAAATCTTTCATTTGAAC 1110
QY 279 AspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSer 298
    |||
Db 1111 GACCAAGTAACTCTCTCAAAATATGAGATCCACGAGATCATTTACACAAATGCTGGCCCTCC 1170
QY 299 LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe 318
    |||
Db 1171 TTGATGATTAAGAATAGGGGTCTCTCATATCCAGGGCCCAAGGCTTATGACAAAGGAGTTT 1230
QY 319 LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal 338
    |||
Db 1231 CTAAAGAGCCTTCGAAACCTTTTGTGACTTTATGAGACCCCAAGTTTGAGTTTCTGCTG 1290
QY 339 LysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIle 358
    |||
Db 1291 AAGTTCATGCACTGGAAATGATGACAGCAGCTTGGCAATAATTTATGCTCTCATTTATT 1350
QY 359 LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsn 378
    |||
Db 1351 CTCATGTGAGACCGCCGACGTTTGTCTGAATGTGAAGCCCATTTGAAGACATTTACAGACAC 1410
QY 379 LeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerGlnLeuPhe 398
    |||
Db 1411 CTGCTACAAAGCCCTGGAGCTCCAGCTGAAGTGAACCAACCTGAGTCTCCACAGCTGTTT 1470
QY 399 AlAluLysLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 418
    |||
Db 1471 GCCAAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGTCACGGAACACCTGCAGCTA 1530
QY 419 LeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIle 438
    |||
Db 1531 CTGCAGGTGATCAAAAGACGAGACAGACATGACTTTCACCCCTCTGACAGAGATC 1590
QY 439 TyrLysAspLeuTyr 443
    |||
Db 1591 TACAAGGACTGTGTAC 1605

RESULT 4
AAT35334
ID AAT35334 standard; DNA; 1647 BP.
XX
AC AAT35334;
XX
DT 14-APR-1997 (first entry)
XX
DE Peroxisome proliferator activated receptor gamma2 coding sequence.
XX
KW Peroxisome proliferator activated receptor gamma; hPPARgamma; human;
KW peroxisome; cholesterol metabolism; lipid metabolism; respiration;
KW fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;
KW herbicide; hPPARgamma2; adipose tissue; obesity; diabetes; anorexia;
KW hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;
KW metabolic disorder; therapy; ss.
XX
OS Homo sapiens.
XX
PH Location/Qualifiers
FT Key 130..1647
FT CDS /tag= a
FT /product= hPPARgamma2 subtype 1
FT /tag= b
FT /product= hPPARgamma2 subtype 2
FT CDS 220..1647
FT /tag= c

```

```

FT /product= hPPARgamma2 subtype 3
XX
XX WO9623884-A2.
XX
PD 08-AUG-1996.
XX
XX 29-JAN-1996; 96WO-US01469.
XX
XX 23-OCT-1995; 95US-0005809.
XX 30-JAN-1995; 95US-0380051.
PR 07-JUN-1995; 95US-0484487.
XX
PA (LIGAND) LIGAND PHARM INC.
XX
PI Mukherjee R;
XX
XX WPI: 1996-384114/38.
DR P-PSDB; AAR93326-R99328.
XX
PT Human peroxisome proliferator activated receptor(s), hPPAR-gamma and
PT hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat
PT e.g. obesity, anorexia and diabetes
XX
XX Claim 43; Page 87-88; 108pp; English.
XX
CC This sequence represents the coding sequence for the human peroxisome
CC proliferator activated receptor gamma2 (hPPARgamma2). This sequence
CC encodes three different hPPARgamma2 subtypes, which differ in their
CC N-terminal sequences. Peroxisomes contain enzymes for cholesterol and
CC lipid metabolism, and respiration. Peroxisome proliferators increase the
CC capacity of the peroxisomes to metabolise fatty acids, via increased
CC expression of the beta-oxidation cycle enzymes. Peroxisome proliferators
CC include unsaturated fatty acids, hypolipidemic drugs, and herbicides.
CC hPPARgamma has two main subtypes (hPPARgamma, and hPPARgamma2), which
CC differ in their N-terminal sequences, and are expressed at high levels in
CC adipose tissue. hPPARgamma proteins repress hPPARalpha activity. The
CC encoded proteins can be used in a cell system with a reporter gene (and
CC optionally a hPPARalpha protein) to identify agonists and antagonists of
CC the polypeptides, which can be used as therapeutic agents. The agonists
CC and antagonists can be used to treat obesity, diabetes, anorexia,
CC hyperlipidaemia, hypercholesterolaemia, hyperlipoproteinaemia, and other
CC metabolic disorders. This sequence, the encoded proteins, and antibodies
CC against them, can be used to establish a tissue specific expression
CC pattern for hPPARgamma.
XX
SQ Sequence 1647 BP; 472 A; 389 C; 371 G; 415 T; 0 other:

Alignment Scores:
Pred. No.: 1.93e-197 Length: 1647
Score: 2231.00 Matches: 443
Percent Similarity: 87.72% Conservative: 0
Best Local Similarity: 87.72% Mismatches: 0
Query Match: 63.60% Indels: 62
DB: 17 Gaps: 1

US-09-931-007A-1 (1-688) x AAT35334 (1-1647)
QY 1 MetGlyGluThrLeuGlyAspSerProIleAspProGluSerAspSerPheThrAspThr 20
    |||
Db 130 ATGGGTGAACCTGGGAGATTCCTATTGACCCAGAAACGATTCCTCAGCTGATACA 189
QY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrPro 40
    |||
Db 190 CTGCTGCAAAACATATTCACAGAAATGACATGTTGTTGACACAGAGATGCCATTCGGCC 249
QY 41 ThrAsnPheGlyIleSerSerValAspPheSerValMetGluAspHisSerHisSerPhe 60
    |||
Db 250 ACCAAGCTTTGGAGACAGCTCCGCGATCTCCGTAATGAAGAACATCCCATCCCTTT 309
QY 61 AspIleLysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAsp 80
    |||
Db 310 GATATCAAGCCCTTCACACTGATGACTTTCACAGCATTTCTACTCCACATTAAGAGAC 369

```

Oy	81	IleProPheThrTrpArgThrAspProValAlaAspTyrIleTyrAspLeuIleuGln	100
Db	370	ATTCATTCCAAACACACAGATCCAGTGGTTCAGATTACAGATATGACCTGAAACTCCAA	429
Oy	101	GluTyrGlnSerAlaIleTysValGlnProAlaSerProCofrTyrIleSerGluTyrThr	120
Db	430	GAGTACCAAAATGCAATCCAAAGTGGACCTGCATCTCCACCTTATTTCTGGAAGACT	489
Oy	121	GlnLeuTyrAsn-----	124
Db	490	CAGCTCTCAATTAAGCCTCATGAGAGCCTTCCAACTCCCTCATGGCAATGAAATGCTG	549
Oy	124	-----	124
Db	550	GTCGTGGAGATTAAGCTTCGTGATTTCACTATGAGATTTCATGCTTTGTAAGATGCAAG	609
Oy	124	-----	124
Db	610	GGTTTCTCCGGACAAACATTCAGATTGAACCTTATCTATGACGATGTATCTTAACGT	669
Oy	125	-----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla	138
Db	670	CGGATCCCAAAAAAGTAGAAATAATGTCATACACTGTCGGTTTCAGAAATGCCCTTGA	729
Oy	139	ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys	158
Db	730	GTGGGAGTCTCAATATGCCATTCAGGTTTGGGCGATGGCACAGGCCAGAGAGAGAG	789
Oy	159	LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg	178
Db	790	CTGTGGCGGAGATCTCCAGTGAATGACACTGATATCCAGAGTCCGCTGACTCCGG	849
Oy	179	AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLys	198
Db	850	GCCCTGGCAAAACATTGTGTATGATCATCATATCAATAGCTCTCCGCTGACCAAGCAAG	909
Oy	199	AlaArgAlaIleLeuThrGlyLysThrThAspLysSerProPheValIleTyrAspMet	218
Db	910	GCGAGCGGATCTTGACAGAGAAACACAAACAGCAAAATCACCATCTCTTATCATGACATG	969
Oy	219	AsnSerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlu	238
Db	970	AATTCCTTAATGATGGAGAGAAATTAATTAAGTTCAAAACACTCACCCCTGCAAGAG	1029
Oy	239	GlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAla	258
Db	1030	CAGAGCAAAAGAGTGGCCATCCGATCTTCAAGGCTGCCAAGTTTCCGCTGAGAGCT	1089
Oy	259	ValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsn	278
Db	1090	GTGCAGGAGATCACAGATATGCCCAAAAGCATTCCGTGGTTTGTAAATCTTGACTTGAAC	1149
Oy	279	AspGlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuLysSer	298
Db	1150	GACCAAGTAACCTCTCCCAATATGAGATGCCAGAAATCATTTACCAATAGCTGGCTCC	1209
Oy	299	LeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGluPhe	318
Db	1210	TTGATGATTAAGATGGGGTTCATATCCGAGGCGCAAGGGCTCATGACAGGAGATT	1269
Oy	319	LeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaVal	338
Db	1270	CTAAGAGCCCTGCCAAAGCCCTTTGGTATTTATGAGCCCAAGTTTATGAGTTTCTGTG	1329
Oy	339	LysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIle	358
Db	1330	AAGTTCAATGACACGGAAATTAAGATGACGCGACTTGGCAATATTTATTTGCTCATTTAT	1389
Oy	359	LeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspSn	378
Db	1390	CTCAGTGGAGACGCCCGCAGGTTTGCCTGAATGTAAGACCCATTGAACACTTTGAACACAC	1449
Oy	379	LeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluLysSerGlnLeuPhe	398

Dd	1450	CTGCTACAGCCCTGGAGCTTCACGTGAAGCTGAACACCCTGAAGTCTCAGAGCTGTTT	1509
Oy	399	AAlaYsLeuEnGIInLySmEtTrAsPLeUrArGlIleValThrGlUnHIsValGIInLeu	418
Dd	1510	GCCAAAGCTGCCTCCAGAAAATGACAGACACTCAACAATTGTCAAGCAACACTGCAGCTA	1569
Oy	419	LLeuGIInValIleLysLysThrGIUThrAsPmetSerLeuHIspProLeuEnGIInIule	438
Dd	1570	CTCGAGGTGATCAAGAAGACAGAGACAGACATGATGCTTCACCCGCTCGCAGAGATC	1629
Oy	439	TyrTLySAsPLeUrTyR 443	
Dd	1630	TACAAGGACTTGTAAC 1644	
RESULT 5			
ID	AAH76294		
AC	AAH76294 standard; cDNA; 2596 BP.		
XX			
XX	AAH76294:		
DT	29-OCT-2001 (first entry)		
DE	Human PAX8e9(-exon 8)-PARgammaae1 cDNA sequence.		
XX			
KW	PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;		
RW	follucular carcinoma; PAX8e9(-exon 8)-PPARGammaae1; human; ss.		
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	161..2596	
FT		/*tag= a	
PN	WO200152789-A2.		
PX	26-JUL-2001.		
PD			
XX			
PX	18-JAN-2001; 2001WO-USO1664.		
XX			
PR	20-JAN-2000; 2000US-0177109.		
BR	14-AUG-2000; 2000US-0225079.		
XX			
PA	(BGM ) BRIGHAM & WOMENS HOSPITAL INC.		
PI	Kroll TG, Fletcher JA;		
XX			
DR	WPI; 2001-514487/56.		
DR	P-PsDB; AAB85801.		
PT	New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and		
PT	treating certain tumors or cancers, e.g. follicular carcinoma -		
XX			
PS	Claim 1; Page 127-131, 145pp; English.		
XX			
CC	The invention relates to an oncogene designated PAX8-PPARGamma1 that		
CC	contains a PAX8 coding region fused to PPARGamma1 coding region. The		
CC	PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant		
CC	methodology. A PPARGamma1 ligand or agent is useful for treating a		
CC	subject having a disorder characterized by the presence of a PAX8-		
CC	PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.		
CC	The PAX8-PPARGamma1 molecules are also useful for providing nucleotide		
CC	and amino acid sequences useful for detecting the above disease. The		
CC	present sequence represents a human PAX8e9(-exon 8)-PPARGammaae1		
CC	polypeptide encoding cDNA.		
SQ	Sequence 2596 BP; 660 A; 745 C; 672 G; 519 T; 0 other;		
Alignment Scores:			
Pred. No.:	1,23e+185	Length:	2596
Score:	2106.50	Matches:	425
Percent Similarity:	84.45%	Conservative:	4
Best Local Similarity:	83.66%	Mismatches:	14



Query Match:	50.05%	Indels:	65
DB:	22	Gaps:	3

US-09-931-007A-1 (1-688) x AAH76294 (1-2596)

Oy	1	MetGIyGurhIreu---	GIyASerProIleAspProGIuSer-----	AspSerPhe	17
Db	1070	GTGGGGCCCAAGCTGCTCCCGGATATCCCAACCCACATCCCAACGAGGAGGAGGCGAT	1129		
Oy	18	ThraSPthIreUSerAlaasniIeSerGIuMetThrMetValAspThrGIuMetPro	37		
Db	1130	GCCTCTCTGCGACATCCGAGGACATGTCGGCAGAAATACCATGTTGCACACAGATGCCA	1189		
Oy	38	PheTrProthrAsnPheGIyIleSerSerValAspLeuSerValMetGIuAspHisSer	57		
Db	1190	TTTCGGCCCAACCACTTGGGATCACCTCCGTGGATCTCTCCGTAAATGAAGACACATCC	1249		
Oy	58	HisSerPheAspIleIySProPheThrTrhValAspPheSerSerIleSerThProHis	77		
Db	1250	CACCTCTTGTATACAGGCCCTTCACTACTGTGATTCCTCAGCATTTCACTCCACAT	1309		
Oy	78	TyrGIuAspIleProPheThrAgtThrAspProValAlaAspTyrIySAspLeu	97		
Db	1310	TACGAGACATTCATTCACACAGACAGATCCAGTGGTTCAGATTAAGATATGACTG	1369		
Oy	98	LysLeuGIuGIuTyrGIuSerAlaIleIySValGIuProAlaSerProProTyrIyS	117		
Db	1370	AAACTCAAGAGTACCAAAAGTCACATCAAGAGGAGCCGCATGCCACTTATATTC	1429		
Oy	118	GIuIySThrGIuLeuTyrAsn-----	124		
Db	1430	GAGAACTACAGCTTACATTAAGCCTCATGAAGACCTTCCACTCCCTCATGGCAAT	1489		
Oy	124	-----	124		
Db	1490	GAATGCTGCTGTGTGAGATTAAGCTTTCGATTTCACTATGAGATTCAATGCTTGTGA	1549		
Oy	124	-----	124		
Db	1550	GGATGCAAGGGTTCTTCGAGACAACATCAGATTGAACCTTATCATGACAGATGAT	1609		
Oy	125	-----ArgasniIySValTyrCysArgPheGIuIyS	135		
Db	1610	CTTAACCTCGGATCCACAAAAAAGTAATAATGATCATCTACTGCGGTTTCAGAA	1669		
Oy	136	CysLeuAlaValGIuMetSerHisAsnAlaIleArgPheGIuArgMetProGIuAlaGIu	155		
Db	1670	TGCCCTGCAGTGGGATGTCTCATTAATGCCATGAGTTTGGGCGATGCCACAGCGAG	1729		
Oy	156	LysGIuIySLeuLeuAlaGIuIleSerSerAspIleAspGIuLeuAsnProGIuSerAla	175		
Db	1730	AAGGAGAAAGCTTTGGCGGAGATCTCCATGATATCGACACACTAATCCAGATCCGCT	1789		
Oy	176	AspLeuArgAlaLeuAlaIySHisLeuTyrAspSerTyrIleIyS	195		
Db	1790	GACCTCCGGCGCCCTGGCAAAACATTTGTATGATCATACATAAAGTCTTCCCGCTGAC	1849		
Oy	196	LysAlaIySAlaArgAlaIleLeuThrGIuIySThrTrhAspIyS	215		
Db	1850	AAACCAAGGCGAGGCGGATCTTGACAGGAAAGACAAACGACAAATCCACCTTGTATC	1909		
Oy	216	TyrAspMetAsnSerLeuMetGIuIySAspIySValIyS	235		
Db	1910	TATACATGATTTCTTAATGATGGGAGAGAAAGTAAATATCAAGTTCAACATACATACCCC	1969		
Oy	236	LeuGIuGIuGIuSerIySValAlaIleArgIlePheGIuIySValIyS	255		
Db	1970	CTGCAGAGACAGACAAAGAGTGGCCATCCCATCTTCAAGGCTCGCATTTCCGCTCC	2029		
Oy	256	ValGIuAlaValGIuGIuIleThrGIuTyrAlaIyS	275		
Db	2030	GTGGAGCTGTGAGGAGATCACAGATGTGCCAAAGCATTTCCGTTTGTAAATCTT	2089		

QY	276	AspleuasnaspGlnValIthrIleuLeuLysTyrGlyValHisGluIleIleTyrThrMet	295
Db	2090	GACTTGAACGCACAGTAACCTCTCCCAATAAGAGAGTCCACAGATCATTTTACCAATG	2149
QY	296	LeuAlasSerLeuMetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThr	315
Db	2150	CTGGCCCTCCTTGATGATTAAGATGGGGTTCTCATATCCGAGGGCCAGGCTTCATGACA	2209
QY	316	ArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMetClnProLysPheGlu	335
Db	2210	AGGGAGTTTCTAAAGAGCCTCGGCAAGACCTTTGGTGACTTTATGAGGCCCAAGTTTGAG	2269
QY	336	PheAlaValLysPheAsnAlaLeuGluLeuAspAspSerAspLeuAlaIlePheIleAla	355
Db	2270	TTTCTGTGAAGTTCAATGCACCTGGAATTAGATGACAGCGACTTGGCAATATTTATTGCT	2329
QY	356	ValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIle	375
Db	2330	GTCAATTATTCATCATGGAGACCCGCCAGGTGGCTGAAATGTAACCCCATTTGAACACATT	2389
QY	376	GlnAspAsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSer	395
Db	2390	CAGACACAACCTCGCTACAGCCCTCGAGCTCCAGCTGAACCTGAACACCCTGAGTCTCA	2449
QY	396	GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHis	415
Db	2450	CAGCTGTTTGGCAGAGCTGCTCCAGAAAATGACAGACCTCAGACAGATTGTCCAGGAACAC	2509
QY	416	ValGlnLeuLeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeu	435
Db	2510	GTGCAGCTACTGCGAGTGATCAAGAAAGAGGAGACAGACAGCATGATGCTTACCCGCTCTG	2569
QY	436	GlnGluIleTyrLysAspLeuTyr	443
Db	2570	CAGAGATCTACAGAGCTGTGAC	2593
RESULT 6			
AAH76283	standard; cDNA; 2625 BP.		
AC	AAH76283;		
XX	29-OCT-2001 (first entry)		
DT	XX		
XX	XX		
DE	Human PAX8e9-PPARGammae1 cDNA sequence.		
XX	XX		
KW	PAX8-PPARGamma1; oncogene; cytostatic; PAX8; PPARGamma1; cancer;		
KW	follicular carcinoma; PAX8e9-PPARGammae1; human; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
FH	Key	Location/Qualifiers	
FT	CDS	1..2625	
FT	/*tag= a		
XX	MO200152789-A2.		
PN	XX		
PD	26-JUL-2001.		
XX	XX		
PF	18-JAN-2001; 2001MO-US01664.		
XX	XX		
PR	20-JAN-2000; 2000OUS-0177109.		
PR	14-AUG-2000; 2000OUS-0225079.		
XX	XX		
PA	(BGHM ) BRIGHAM & WOMENS HOSPITAL INC.		
XX	XX		
P1	KIO11 YG, Fletcher JA;		
XX	XX		
DR	WPI: 2001-51487/56.		
DR	P-PSDB: AAB85795.		
XX	XX		
PT	New PAX8-PPARGc1 oncogene and oncoprotein, useful for detecting and treating certain tumors or cancers, e.g. follicular carcinoma -		

XX Claim 1: Page 112-116; 145bp; English.  
PS  
XX The invention relates to an oncogene designated PAX8-PPARgamma1 that  
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The  
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant  
CC methodology. A PPARgamma1 ligand or agent is useful for treating a  
CC subject having a disorder characterized by the presence of a PAX8-  
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.  
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide  
CC and amino acid sequences useful for detecting the above disease. The  
CC present sequence represents a human PAX8-PPARgamma1 polypeptide  
CC encoding cDNA.  
XX  
SQ Sequence 2625 BP; 662 A; 761 C; 646 G; 556 T; 0 other;  
SQ  
Alignment Scores:  
Pred. No.: 1,25e-185 Length: 2625  
Score: 2106.50 Matches: 425  
Percent Similarity: 84.45% Conservative: 4  
Best Local Similarity: 83.66% Mismatches: 14  
Query Match: 60.05% Indels: 65  
DB: Gaps: 3  
US-09-931-007a-1 (1-688) x AAH76283 (1-2625)  
QY 1 MetGlyGlnThrLeu---GlyAspSerProIleAspProGlnSer-----AspSerPhe 17  
Db 1099 GTGGGGCCCGCCGCGCCGATACCCACACATCCCGCCACGAGGACGCGAGCTAT 1158  
QY 18 ThrAspThrLeuSerAlaAsnIleSerGlnGluMetThrMetValAlaSprThrGluMetPro 37  
Db 1159 GCCTCCCTCCATCCGAGCAGCATGCTGCAGAAATGCACATGCTTGACACAGAGATGCCA 1218  
QY 38 PheThrProThrAsnPheGlyIleSerSerValAspLeuSerValMetGluAspHisSer 57  
Db 1219 TTCTGGCCCAACCACTTGGATCAGCTCCGATCTCTCCGTAATGGAAGACCACTCC 1278  
QY 58 HisSerPheAspIleLysProPheThrThrValAspPheSerSerIleSerThrProHis 77  
Db 1279 CACTCCCTTGATATCAAGCCCTTCACACTGTGATCTTCCACGATTTCTACTCCACAT 1338  
QY 78 TyrGluAspIleProPheThrArGThrAspProValAlaAlaSprThrLysTyrAspLeu 97  
Db 1339 TAGGAGACATTCATTCACAAAGACAGATCCAGTGGTGCAGATTCACAGATAGACCTG 1398  
QY 98 LysLeuGlnGlnLysThrGlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSer 117  
Db 1399 AAACCTCAAGAGTACCAAAAGTGAATCAAAAGTGAAGCTTCATCTCCACTTATTATTC 1458  
QY 118 GlnLysThrGlnLeuTyrAsn----- 124  
Db 1459 GAAAGAGCTAGCTTACATTAAGCCATGAAGAGCCTTCCAACCTCCATATGGCAATT 1518  
QY 124 ----- 124  
Db 1519 GAATGTCGTCTGTGAGAGATAAAGCTTCTGATTCATATGAGATTACGCTTGCAA 1578  
QY 124 ----- 124  
Db 1579 GGATGCAGAGGTTCTTCCGAGAACATCAGATTGAAGCTTATCATAGACAGATGTAT 1638  
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLys 135  
Db 1639 CTTAAGTGTGATCCACAAAAAAGTAGAATTAATGTAGTACTGTGCTTTCAGAAA 1698  
QY 136 CysLeuAlaValAlaLysMetSerHisAsnAlaIleArgPheGlnLysArgMetProGlnAlaGlu 155  
Db 1699 TGCCTTGCAATGGGAGTGTCTCATATGACATCAGGTTTGGCGGAGTGCACAGGCGCAG 1758  
QY 156 LysGlnLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAla 175  
Db 1759 AAGGAGAAGCTGTGGCGGAGATCTCCAGTATATGACACAGCTGAATCCAGAGTCCGCT 1818

QY 176 AspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr 195  
Db 1819 GACCTCCGGCCCTCGGCAAAACATTTTATGACTCATACATAAGTCTCTCCGCTGACC 1878  
QY 196 LysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIle 215  
Db 1879 AAACCAAGCGGAGCGGATCTTGACAGAAAGACACACAAATCACCATTCCGTTATC 1938  
QY 216 TyrAspMetLysSerLeuMetMetGlyGlnAspLysIleLysPheLysHisIleThrPro 235  
Db 1939 TATGACATGTAATCTTAAATGATGGGAGAGATAAAATCAAGTTCAAAACATCACCCCC 1998  
QY 236 LeuGlnGlnLysSerLysGlnValAlaIleArgIlePheGlnGlnLysGlnPheArgSer 255  
Db 1999 CTCGAGAGCAGACAAAGAGCTGCGCATCCGATCTTTAGGGCTCCGAGTTCCGCTCC 2058  
QY 256 ValGlnAlaValGlnGlnLysIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeu 275  
Db 2059 GTGGAGGCTGTGCAGGAGATCAGAGATATGCCAAAGCATCTCGGTTTGTAAATCTT 2118  
QY 276 AspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMet 295  
Db 2119 GACTTGACAGCAGCAAGTAACCTCTCAATATGGAATGGAATCCACAGATCATTTACAAATG 2178  
QY 296 LeuAlaSerLeuMetAlaAsnLysAspGlyValLeuIleSerGlnGlnGlnLysPheMetThr 315  
Db 2179 CTGGCCCTCTGATGATTAAGATGGGTTCTCATATCCAGGAGCCCAAGCTTCATGACA 2238  
QY 316 ArgGlnPheLeuLysSerLeuArgLysProPheGlyAspPheMetGlnProLysPheGln 335  
Db 2239 AGGAGATTTCTAAAGAGCCCTGCAGAAAGCCCTTGGTCACTTATGAGGCCCAAGTTTGAG 2298  
QY 336 PheAlaValLysPheAsnAlaLeuGlnLeuAspAspSerLysPheLeuAlaIlePheIleAla 355  
Db 2299 TTTCGCTGAAGTTCAATGCACTGGAATTAGATGACAGCCGATTTGCCAATATTTATTCCT 2358  
QY 356 ValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIle 375  
Db 2359 GTCATATTCTTCACTGTGAGAGCCGCCAGCTTCTGATATGTAAGCCATTTGAAGACATT 2418  
QY 376 GlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnSerSer 395  
Db 2419 CAAGACAACTGCTTACAGGCCCTGAGACCTCCAGCTGAAGCTGAACCAACCTGAGTCTCA 2478  
QY 396 GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHis 415  
Db 2479 CAGCTGTTTGCAGAGCTGCTCCAGAAATGACAGACCTCAGACAGATTTGCACGGAACAC 2538  
QY 416 ValGlnLeuLeuGlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeu 435  
Db 2539 GTGCAGCTACTGACAGTGATCAAGAGACGAGACGACAGACATGATGCTTCCACCGCTCTC 2598  
QY 436 GlnGlnLysTyrLysAspLeuTyr 443  
Db 2599 CAGGAGATCTACAAAGGACTGTGAC 2622  
RESULT 7  
AAH76281  
ID AAH76281 standard; cDNA; 2334 BP.  
XX  
AC AAH76281;  
XX  
DT 29-OCT-2001 (first entry)  
XX  
DE Human PAX8-PPARgamma1 cDNA sequence.  
XX  
XX PAX8-PPARgamma1; oncogene; cytosolic; PAX8; PPARgamma1; cancer;  
KW follicular carcinoma; PAX8-PPARgamma1; human; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers

FT CDS 1..2334  
FT /\*tag= a  
XX MO200152789-A2.  
XX 26-JUL-2001.  
XX 18-JAN-2001; 2001WO-US01664.  
XX 20-JAN-2000; 2000US-0177109.  
XX 14-AUG-2000; 2000US-0225079.  
XX (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.  
XX PA  
XX Krcoll TG, Fletcher JA:  
XX WPI: 2001-514487/56.  
XX DR P-PSDB: AAB85793.  
XX  
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and  
XX treating certain tumors or cancers, e.g. follicular carcinoma -  
XX  
XX Claim 1: Page 100-104; 145pp; English.  
XX  
XX The invention relates to an oncogene designated PAX8-PPARgamma1 that  
XX contains a PAX8 coding region fused to PPARgamma1 coding region. The  
XX PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant  
XX methodology. A PPARgamma1 ligand or agent is useful for treating a  
XX subject having a disorder characterized by the presence of a PAX8-  
XX PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.  
XX The PAX8-PPARgamma1 molecules are also useful for providing nucleotide  
XX and amino acid sequences useful for detecting the above disease. The  
XX present sequence represents a human PAX8e7-PPARgamma1 polypeptide  
XX encoding cDNA.  
XX  
XX SO Sequence 2334 BP: 614 A; 650 C; 573 G; 497 T; 0 other:  
  
Alignment Scores:  
Pred. No.: 5,34e-185 Length: 2334  
Score: 2099.00 Matches: 420  
Percent Similarity: 84.33% Conservative: 5  
Best Local Similarity: 83.33% Mismatches: 13  
Query Match: 59.83% Indels: 66  
DB: 22 Gaps: 2  
  
US-09-931-007a-1 (1-688) x AAB76281 (1-2334)  
OY 6 G1yAsPSePrOIleAsPProGluSeRaSPSeRheThrAsPThLeuSeRa1AAsn--- 24  
DB 820 GGGAGGCGCACCCCTGACCTCCACACGCGCAGTGGGCGCAGACCTCTCGACCTCACAG 879  
OY 25 -----IleSeRGlInuMeThrMeTValAsPThRInGluMeProPhetProthr 41  
DB 880 ACCTACCCCGTGGTGGCAAAATGACATGTGTGACACAGAGATCCATTCTGGGCCAC 939  
OY 42 AsnPhGlyIleSeSeSerValAsPSeSerValMeTGlAsPHisSeRHisSeRPhSeAsP 61  
DB 940 AACTTGGGATAGCTCCGTGGATCTCTCCGTAATGAAGAACCACTCCACCTCTTGAT 999  
OY 62 IleLySProPhetThrThValAsPSeSeSerIleSeSerThProHisTyrGlInuSpIle 81  
DB 1000 ATCAAGCCCTTCACACTGCTGACTTCTCCAGCATTTCTACTCCACATTAAGAACATT 1059  
OY 82 ProPhetThrArgThAsPProGluValAlaAsPTrLysTyrAsPLeuLysInGlu 101  
DB 1060 CCATTTCACAGAACAGATCCAGTGTGACAGATTACAGATAGACCTGAAACTTCAAGAG 1119  
OY 102 TyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluInuThRIn 121  
DB 1120 TACCAAAAGTGCATCAAGATGAGGCTGCATCTCCACCTTATATTCTGAGAGACTGAG 1179  
OY 122 LeuTyrAsn----- 124  
|||||

DB 1180 CTCTACATAAGCCCTCATGAAGAGCCTTCCAACTCCCTCATGGCAATTGAATGTCGTCTC 1239  
OY 124 ----- 124  
DB 1240 TGTGAGATAAAGCTTCTGGATTTCTACTATGAGATTCTGCTTGTGAAGATGCAAGGCT 1299  
OY 124 ----- 124  
DB 1300 TTCTCCGAGAACATCAGATTGAAGCTTATCTATGACAGATGATCTTAACGTGCG 1359  
OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaVal 139  
DB 1360 ATCCACAAAAAAGTAGAAATAAATGTCTACTGTCTCGTTTCAGAAATGGCTTGCACTG 1419  
OY 140 GlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLysLeu 159  
DB 1420 GGGATGCTCATTAATGCCATCAGGTTTGGGGGATGCCACAGGCCGAGCAAGAGAGAGCTG 1479  
OY 160 LeuAlaGluIleSeSerAsP1leAsPInLeuAsnProGluSerAlaAsPLeuArgAla 179  
DB 1480 TTGGCGGAGATCTCCAGTATATGACAGCTGAAATCCAGAGTCCGCTGACCTCGGGCC 1539  
OY 180 LeuAlaLysHisLeuTyrAsPSeTyrIleLysSerPheProLeuThrLysAlaLysAla 199  
DB 1540 CTGGCAAAACATTTGTATGACTCATACATAAAGTCTTCCGCTGACCAAAAGAGCG 1599  
OY 200 ArgAlaIleLeuThrGlyLysThrThrAsPLeuSerProPheValIleTyrAsPMetLsn 219  
DB 1600 AGGGGATCTTGACGAGGAAGCAACACACAAATCACCATTCCTGTATATATGACATGAT 1659  
OY 220 SerLeuMetMetGlyLysAsPLeuLysIleLysPheLysHisIleThrProLeuGlnGluIn 239  
DB 1660 TCTCTAATGATGGGAGGAATTAATCAAGTTCAACACATCACCCCTCGAGAGAGCAG 1719  
OY 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259  
DB 1720 AGCAAAAGAGTGGCCATCCGATCTTTCAGGGCTGCGCATTTCCGCTCGTGGAGCTGTG 1779  
OY 260 GlnGluIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAsnLsn 279  
DB 1780 CAGGAGATCACAGATATGACCAAAAGCATCTCTGTTTGTAAATCTTGACTTGAAAGAC 1839  
OY 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299  
DB 1840 CAAGTAACCTCCCTCAATATGAGACTCCAGCAGATCATTTACACATCTGGCTCCCTTG 1899  
OY 300 MetAsnLysAspGlyValLeuIleSerGlnGlyInuGlyPheMetThrArgGluPheLeu 319  
DB 1900 ATGAATAAAGATGGGGTCTCATATCCAGAGGCCAAGGCTTCATGACAAGGAGTGTCTA 1959  
OY 320 LysSerLeuArgLysProPheGlyAsPheMetGluProLysPheGluPheAlaValLys 339  
DB 1960 AAGACCTCGCAAAAGCTTTTGGTACTTTATGAGGCCCAAGTTGAATGTGCTGTGAAG 2019  
OY 340 PheAsnAlaLeuGluLeuAsPSeSerAsPLeuAlaIlePheIleAlaValIleLeu 359  
DB 2020 TTCAATGACACTGGAATTGATGACAGCAGCATTGGCAATATTTATTTGCTGCTATTATTC 2079  
OY 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAsP1leGlnAsPAsnLeu 379  
DB 2080 AGTGAGAGACCCCGGAGTTTCTGTAATGTGAAGCCCATTTGAAGACATTCAGACAACTG 2139  
OY 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerGlnLeuPheAla 399  
DB 2140 CTACAGCCCTTGAGCTCAGCTGAAGCTGAACCACTCTGAGTCTCCACACTGTTTGC 2199  
OY 400 LysLeuLeuGlnLysMetThrAsPLeuArgGlnIleValIleThrGlnHisValGlnLeu 419  
DB 2200 AAGCGCTCCAGAAATATACAGACCTCAGACAGATGTGCAGGAACAGCTGACGCTACTG 2259  
OY 420 GlnValIleLysLysThrGluThrAsPMetSerLeuHisProLeuLeuGlnIleLys 439  
DB 2260 CAGGTGATCAAGAAAGACGAGACAGACATGAGTCTTACCCGCTCTGCGAGAGATCTAC 2319  
|||||

QY 440 LysAspLeuTyr 443  
 |||  
 Db 2320 AAGGACTTGTAC 2331

RESULT 8  
 AAT35333

ID AAT35333 standard; DNA; 1936 BP.  
 AC AAT35333;  
 XX  
 XX  
 DT 14-APR-1997 (first entry)  
 XX  
 DE Peroxisome proliferator activated receptor gamma coding sequence.  
 XX  
 KM Peroxisome proliferator activated receptor gamma; hPPARgamma; human;  
 KM peroxisome; cholesterol metabolism; lipid metabolism; respiration;  
 KM fatty acid metabolism; beta-oxidation cycle enzyme; hypolipidemic drug;  
 KM herbicide; hPPARgamma2; adipose tissue; obesity; diabetes; anorexia;  
 KM hyperlipidaemia; hypercholesterolaemia; hyperlipoproteinaemia; antibody;  
 KM metabolic disorder; therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 157..1641  
 FT /\*tag= a  
 FT /product= hPPARgamma subtype 1  
 FT CDS 208..1641  
 FT /\*tag= b  
 FT /product= hPPARgamma subtype 2  
 FT 214..1641  
 FT /\*tag= c  
 FT /product= hPPARgamma subtype 3  
 XX  
 PN MO9623884-A2.  
 XX  
 PD 08-AUG-1996.  
 XX  
 PF 29-JAN-1996; 96MO-US01469.  
 XX  
 PR 23-OCT-1995; 95US-0005809.  
 PR 30-JAN-1995; 95US-0380051.  
 PR 07-JUN-1995; 95US-0484487.  
 XX  
 PA (LIGA-) LIGAND PHARM INC.  
 XX  
 PI Mukherjee R;  
 DR WPI: 1996-384114/38.  
 DR P-PSDB: AAR99323-R99325.  
 XX  
 PT Human peroxisome proliferator activated receptor(s), hPPAR-gamma and  
 PT hPPAR-gamma2 - useful to identify polypeptide (ant)agonists to treat  
 PT e.g. obesity, anorexia and diabetes  
 XX  
 PS Claim 2: Page 83-84; 108pp; English.

This sequence represents the coding sequence for the human peroxisome  
 proliferator activated receptor gamma (hPPARgamma). This sequence  
 encodes three different hPPARgamma subtypes, which differ in their  
 N-terminal sequences. Peroxisomes contain enzymes for cholesterol and  
 lipid metabolism, and respiration. Peroxisome proliferators increase the  
 capacity of the peroxisomes to metabolise fatty acids, via increased  
 expression of the beta-oxidation cycle enzymes. Peroxisome proliferators  
 include unsaturated fatty acids, hypolipidemic drugs, and herbicides.  
 hPPARgamma has two main subtypes (hPPARgamma, and hPPARgamma2), which  
 differ in their N-terminal sequences, and are expressed at high levels in  
 adipose tissue. hPPARgamma proteins repress hPPARalpha activity. The  
 encoded proteins can be used in a cell system with a reporter gene (and  
 optionally a hPPARalpha protein) to identify agonists and antagonists of  
 the polypeptides, which can be used as therapeutic agents. The agonists  
 and antagonists can be used to treat obesity, diabetes, anorexia,

CC hyperlipidaemia, hypercholesterolaemia, hyperlipoproteinaemia, and other  
 CC metabolic disorders. This sequence, the encoded proteins, and antibodies  
 CC against them, can be used to establish a tissue specific expression  
 CC pattern for hPPARgamma.  
 XX

SQ Sequence 1936 BP; 545 A; 483 C; 423 G; 485 T; 0 other;

#### Alignment Scores:

Pred. No.:	5,2e-185	Length:	1936
Score:	2098.00	Matches:	421
Percent Similarity:	82.65%	Conservative:	3
Best Local Similarity:	82.07%	Mismatches:	5
Query Match:	59.81%	Indels:	84
DB:	17	Gaps:	2

US-09-931-007a-1 (1-688) x AAT35333 (1-1936)

QY	28	GlumetThmElvaAlasPhrGlumetProPhetPrProThAspNheGlyIleSerSer	47
Db	205	GAATGACCATGATGTCACACAGATGCCATCTGCGCCACCACTTGGATGCAGCTCC	264
QY	48	ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPhetThr	67
Db	265	GTGGATCTCTCCGTAAATGAGAACACACTCCCTCTTGATATATAGCCCTTCACTACT	324
QY	68	ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPhetThrArgThrAsp	87
Db	325	GTGGACTTCTCCAGCATTTCTACTCCACATTCAGAAACATTCATTCACAAAGACAGAT	384
QY	88	ProValValAlaAspTyrIleTyrAspLeuLysLeuGlnGlyTyrGlnSerAlaIleLys	107
Db	385	CCAGTGGTTCAGAAATTCACAGATGACCGAAACCTTCAGATACCAAGTCACATCAAA	444
QY	108	ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn-----	124
Db	445	GTGGAGCTGCATCTCCACCTTATTTCTGAGAGACTCAGCTTACATATAGCCTCAT	504
QY	124	-----	124
Db	505	GAAGAGCTTCCCACTCCCTCATGCGCAATTCGATGTCGTGTGAGATTAAGCTTCT	564
QY	124	-----	124
Db	565	GGATTTCACTATGAGATTCATGCTTGTGAAGATGCAAGGTTTCTTCCGAGAACATC	624
QY	125	-----Arg 125	684
Db	625	ACATTGAAGCTTATCTATGACAGATGTCATTAACTGTCGATCCACAAAAAAGTACA	684
QY	126	AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla	145
Db	685	AATAATGTCACACTACTGTCGTTTCAGAAATCCCTTCGAGTGGGATGTCATATATGCC	744
QY	146	IleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSer	165
Db	745	ATCAGGTTTGGCGGATGCCACAGCCCAAGAGGAGGAGCTTTCGCGAGTTCCTCCAGT	804
QY	166	AspIleAspGlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyr	185
Db	805	GATATCGACACCTTAATCCAGATCCGTCGACCTCGCGGCTTGCAAAACATTTGAT	864
QY	186	AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly	205
Db	865	GACTCATATCAATAAATCCCTCCGCTGACCAAGCAAGGAGGCGCATTTGCACAGA	924
QY	206	LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetLeuGlu	225
Db	925	AAGACACAGACAAATTCACCATTCCTATCTATGACATCAATTCCTTAATGATGGAGAA	984
QY	226	AspLysIleLysPheLysHisIleThrProLeuGlnGlnGlnSerLysGluValAlaIle	245
Db	985	GATAAATCAAGTTCAACACATCACCCCTTCGAGGAGCAGAGCAAGAGAGGTGCGCATC	1044

```

QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGlyIur 265
    |||
Db 1045 CCACACTTTTCAGAGGCTGGCAGATTTTCCTCCCTGCGAGGCTGTGCAGAGATTCACAGATAT 1104
QY 266 AlaIysSerIleProGlyPheValAsnLeuAspLeuAsnAspGluValThrLeuLeuLys 285
    |||
Db 1105 GCCAAAGCATTCCTGGTTTGTAAATCTTGATCTGAACGACCAAGTAACTCTCTCAAA 1164
QY 286 TTYGIVAlHISGluIleIleIleThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
    |||
Db 1165 TATGAGGTCACGAGATCATTTTACCAATGCTGGCTCTCTTATGAAATTAAGATGGGTT 1224
QY 306 LeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
    |||
Db 1225 CTCATATCCGAGGCGCCAGGCTTCATGACAGAGGTTTCTAAAGACCTGGGAAAGCCT 1284
QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
    |||
Db 1285 TTYTGCTGACTTATGAGAGCCCAAGTTTGAGTTTGCTGTGAAATGCAATGCACTGGAATTA 1344
QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
    |||
Db 1345 GATGACAGGCACTGGCAATATTATGCTGATCATTTCTAGTGAAGACCCGACAGGT 1404
QY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385
    |||
Db 1405 TTGCTGAATGTGAAGCCCTTGAAGCATTCAGACAACTGCTACAAACCCCTGGAGCTC 1464
QY 386 GlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuGlnLysMet 405
    |||
Db 1465 CAGCTGAACTGACACCCCTGAGTCTCCACAGCTGTTTGCCAAAGCTGCTCCAGAAAATG 1524
QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuGlnValIleLysLysThr 425
    |||
Db 1525 ACAGACTCTGACAGATTGTCTCAGGACACGTCGACACTCTCAGGTGATCAAGAAAGACG 1584
QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleLysLysAspLeuTy 443
    |||
Db 1585 GAGACAGACATGAGTCTTACCCGCTCGTCAAGAGATCTACAGAGACTTGTACTAGCAG 1644
QY 443 ----- 443
Db 1645 AGAGTCTGAGCCACTGCCACATTTCCCTCTCCAGTTGCACATATCTGAGCCGGAAT 1704
QY 444 AlaTrrPAlaIleLeuThrLysThrAspLysSer 456
    |||
Db 1705 TCTTTTGTCTTTTACCCTGGAAGAAATCTCATTAAGC 1743

```

```

RESULT 9
AAH76282
ID AAH76282 standard; cDNA; 2523 BP.
AC AAH76282;
XX
XX 29-OCT-2001 (first entry)
XX
XX Human PAX8e8-PPARGammae1 cDNA sequence.
DE
XX
XX PAX8-PPARGammae1; oncogene; cytosolic; PAX8; PPARGammae1; cancer;
KM
XX Folicular carcinoma; PAX8e8-PPARGammae1; human; ss.
KW
OS Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH CDS 1..2523
FT /tag= a
XX
XX WO200152789-A2.
XX
XX 26-JUL-2001.
XX
XX 18-JAN-2001; 2001WO-US01664.
XX

```

```

PR 20-JAN-2000; 2000US-0177109.
XX 14-AUG-2000; 2000US-0225079.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
XX Kroll TG, Fletcher JA;
XX
XX WPI: 2001-514487/56.
DR P-PSDB; AAB85794.
XX
XX New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma
XX
PS Claim 1; Page 106-109; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARGamma1 that
CC contains a PAX8 coding region fused to PPARGamma1 coding region. The
CC PAX8-PPARGamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARGamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARGamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARGamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PAX8e8-PPARGammae1 polypeptide
CC encoding cDNA.
XX
SQ Sequence 2523 BP; 643 A; 723 C; 614 G; 543 T; 0 other:

```

```

Alignment Scores:
Pred. No.: 1,72e-184 Length: 2523
Score: 2094.00 Matches: 416
Percent Similarity: 87.03% Conservative: 0
Best Local Similarity: 87.03% Mismatches: 62
Query Match: 59.69% Indels: 1
DB: 22 Gaps: 1

```

```

US-09-931-007A-1 (1-688) x AAH76282 (1-2523)
QY 28 GluMetThrMetValAspThrGluMetProPheThrAsnPheGlyIleSerSer 47
    |||
Db 1087 GAATGACCATGATGTGACACAGATGCAATTCCTGAGCCACCACTTGGATCAGCTCC 1146
QY 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleLysProPheThrThr 67
    |||
Db 1147 GTGATCTCTCCGTATGAGAAACACCATCCACTCTTGATATCAAGCCCTTCACTACT 1206
QY 68 ValAspPheSerSerIleSerThrProHisTyrgluAspIleProPheThrArgThrAsp 87
    |||
Db 1207 GTTGACTTCTCCAGCATTTCTACTCCACTTCAAGAAATTCATTCACAAGAACAGAT 1266
QY 88 ProValValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLys 107
    |||
Db 1267 CCAGTGTGTCAGATTTACAGATGACCTGAACCTTCAAGATACCAAGTCAATCAAA 1326
QY 108 ValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTy 124
    |||
Db 1327 GTGAGCCGTCATCTCCACTTATTTCTGAGAAGACACACTCTACATATAAGCTCAT 1386
QY 124 ----- 124
Db 1387 GAAGAGCCTTCCAACTCCCTCATGCAATTGAATGTCGTGTGAGATAAAGCTTCT 1446
QY 124 ----- 124
Db 1447 GGATTTCACTATGAGATTCATGTCGAAGATGCAAGGTTTCTTCGGGAAACATTC 1506
QY 125 -----Arg 125
Db 1507 AGATTGAAGCTTATCTATGACAGATGTGATCTTAAGTGTGCGATCCACAAAAGATGCA 1566
QY 126 AsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAla 145
    |||
Db 1567 AATAAATGTCACTACTGTCGCTTCAAGAAATCCCTTGACAGTGGGATGTCTCATATATGCC 1626

```

```
QY 146 IlaargPheGlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSer 165
    |||
DB 1627 ATCAGGTTTGGCGGATGCCAGAGCCGAGAGAGAACTGTTGGCGGAGATCTCCAGT 1686
QY 166 AspIleAspGlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysIleTyr 185
    |||
DB 1687 GATATGACACGAGTGAATCCAGAGTCCGCTGACCTCCGGGCGCTGGCAAAACATTTGTAT 1746
QY 186 AspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
    |||
DB 1747 GACTCATACATAAAGTCTTCCCGCTGACCAAGCAAGGCGAGGCGATCTTGACAGGA 1806
QY 206 LysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetMetGlyGlu 225
    |||
DB 1807 AAGACAACAGACAATCACCATTCGTATCTATGACATGAAATTCCTTAATGATGGAGAA 1866
QY 226 AspLysIleLysPheLysHisIleThrProLeuGlnGluInSerLysGluValAlaIle 245
    |||
DB 1867 GATAAATCAAGTTCAAAACATCACCCCTCCAGAGAGAGCAAAAGAGTGGCCATC 1926
QY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyr 265
    |||
DB 1927 CGCATCTTTCAGGCGCTGCCAGTTCCGCTCGTGAGGCTGTGCAGAGATCACAGAGTAT 1986
QY 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLys 285
    |||
DB 1987 GCCAAAAGCATTCCTGTTTGTAAATCTTGACTGAACAGACCAAGTAATCTCCCAAA 2046
QY 286 TyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeuMetLysAspGlyVal 305
    |||
DB 2047 TATGGATCCACAGATCATTTACAAATGCTGCTCTTGATGAATAAAGTGGGGTT 2106
QY 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325
    |||
DB 2107 CTCATATCGAGGGCCAAAGGCTTCATGACAGAGGAGTTTCTAAAGACCTCGCAAAAGCT 2166
QY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345
    |||
DB 2167 TTTGGTACATTTATGAGAGCCCAAGTTGAGTTGCTGTAAGTTCAATGCACATGGAATTA 2226
QY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlnAspArgProGly 365
    |||
DB 2227 GATGACAGCGACTGGCAAAATTTATTTGCTGCTATTATCTTCAGTGAGAGCCCCAGAGT 2286
QY 366 LeuLeuAsnValLysProIleGlnAspIleGlnAspAsnLeuGlnAlaLeuGluLeu 385
    |||
DB 2287 TTGCTGATATGTGAAGCCCATTTGAAGACATTCAGACAACTGCTACAGAGCCCTGGAGCTC 2346
QY 386 GlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAlaLysLeuGlnLysMet 405
    |||
DB 2347 CACGTGAAGCTGAACCACTGAGTCTCTCACACGCTTTGGCAAGCTGCTCCAGAAATG 2406
QY 406 ThrAspLeuArgGlnIleValThrGluHisValGlnLeuGlnValIleLysLysThr 425
    |||
DB 2407 ACAGACCTCAGACAGATTTGTCAGGAACAGTGCACCTACTGCAAGTGTATCAAGAGACG 2466
QY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyrLysAspLeuTyr 443
    |||
DB 2467 GAGACAGACATGATGTTTCAACCCGCTCTGCAAGAGATCTACAGAGACTTTGTAC 2520
RESULT 10
AAD21021
ID AAD21021 standard; cDNA; 1434 BP.
XX
XX AAD21021;
AC
XX
XX 15-JAN-2002 (first entry)
DE Human peroxisome proliferator activated receptor (PPAR) gamma1 cDNA.
XX
XX Human peroxisome proliferator activated receptor gamma1; PPAR gamma1;
KW cytotoxic; antiproliferative; antiviral; cancer; cell proliferation;
```

```
KW viral infection; pharmaceutical; thiazolidinedione; ss.
XX
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1434
FT /tag= "Human peroxisome proliferator activated
FT /product= receptor (PPAR) gamma1 protein"
FT
PN US6294559-B1.
XX
PD 25-SEP-2001.
XX
PF 03-AUG-1998; 98US-0128142.
XX
PR 02-MAY-1996; 96US-016694P.
PR 18-APR-1997; 97US-0844007.
XX
XX (MER1 ) MERCK & CO INC.
XX
PI Smith RG;
XX
DR WPI: 2001-647265/74.
DR P-PSDB: AAE12867.
XX
XX Use of thiazolidinedione for treating cancer and viral infections -
PS Example 3; Fig 1; 17pp; English.
XX
XX The invention relates to compounds and ligands that bind to human
CC peroxisome proliferator activated receptors (PPAR) gamma1 and gamma2.
CC The invention is useful for treating cancer and other disorders including
CC excessive cell proliferation and viral infection. The invention is also
CC directed to the use of PPAR gamma1 and gamma2 to identify compounds that
CC are antiproliferative, antiviral and antitumour agents. The invention
CC also relates to a method of treating cancer using a pharmaceutical
CC composition comprising thiazolidinedione in an amount sufficient to
CC modulate PPAR gamma1 and gamma2 activity. The present cDNA sequence
CC encodes human peroxisome proliferator activated receptor (PPAR) gamma1
CC protein related to the invention.
XX
SQ Sequence 1434 BP; 411 A; 343 C; 330 G; 350 T; 0 other:
XX
Alignment Scores:
Pred. No.: 2,42e-184 Length: 1434
Score: 2089.00 Matches: 415
Percent Similarity: 87.00% Conservative: 0
Best Local Similarity: 87.00% Mismatches: 0
Query Match: 59.55% Indels: 62
DB: 22 Gaps: 1
US-09-931-007A-1 (1-688) x AAD21021 (1-1434)
QY 29 MetThrMetValAspThrGluMetProPheThrProThrAsnPheGlyIleSerSerVal 48
    |||
DB 1 ATGACCATGTTGACACAGATGATGCCATTCGCGCCACCAACTTGGGATGAGTCCGCG 60
QY 49 AspLeuSerValMetGlnAspHisSerHisSerPheAspIleLysProPheThrVal 68
    |||
DB 61 GATCTCTCCGTATGAGAACCACTCCACCTCTTTGAATATCAACCCCTCACTACTGTT 120
QY 69 AspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrArgThrAspPro 88
    |||
DB 121 GACTTCTCCAGCATTTCTACTCCACATTTAGAGACATTCCTTACCAAGAACACATCA 180
QY 89 ValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysVal 108
    |||
DB 181 GTGGTTGCAGATTACAGATGACCTGAACCTCAAGAGATACCAAGTCAATCAAGTG 240
QY 109 GluProAlaSerProProTyrTyrSerGluLysThrGlnLeuTyrAsn----- 124
    |||
DB 241 GAGCCTGCATCTCCACTTATTTATTTCTGAGAGAGCTCAAGCTTCAAAATTAAGCCTCATGAA 300
```

```

Oy 124 ----- 124
Db 301 GAGCCTTCAACTCCCTCATGATGATGTGCTGTGAGACATAAAGCTTCTGGA 360
Oy 124 ----- 124
Db 361 TTTCACATGAGATTCATGCTGTGTGAAGATGCAAGGGTTTCTTCCGGAGAACATCAGA 420
Oy 125 -----Argasn 126
Db 421 TTGAAGCTTATCTATGACAGATGTGATCTTAACCTGCGATCCACAAAAGTNGAAT 480
Oy 127 LysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIle 146
Db 481 AAATCTCAGTACTGTGCGGTTTCAGAAATGCCCTTGACAGTGGGATCTTCATATCCCATC 540
Oy 147 ArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeuAlaGlnIleSerSerAsp 166
Db 541 AGGTTTGGGCGGATGCCACAGGCCGAGAGGAGAGCTGTGGCGAGATCTCCAGTGAT 600
Oy 167 IleAspGlnLeuAsnProGlnSerIleAspLeuArgAlaLeuAlaLysIleuTyrAsp 186
Db 601 ATCGACCAAGCTGAATCCAGAGTCCGCTGACCTCCGGGCCCTGGCAAAACATTGTATATAC 660
Oy 187 SerTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLys 206
Db 661 TCATACATTAAGTCTCTCCGCTGACCAAGCAAGGCGAGGCGATCTTGACAGAGAAAG 720
Oy 207 ThrThrAspLysSerPheProPheValIleTyrAspMetAsnSerLeuMetGlyLysAsp 226
Db 721 ACAACGAGCAAAATCACCATTCGTATCTATGACATGAATTCCTTATGATGGAGAAATAT 780
Oy 227 LysIleLysPheLysHisIleThrProLeuGlnGlnIleSerLysGlnValAlaIleArg 246
Db 781 AAAATCAAGTTCAACACATCACACCCCTGCGAGGAGCAGACCAAGAGAGTGGCCATCCGC 840
Oy 247 IlePheGlnLysCysGlnPheArgSerValGlnAlaValGlnGlnIleThrGlnTyrAla 266
Db 841 ATCTTTCAGGGCTGCGCAATTCCTGCGCTGCGAGGCTGTGCAGAGATCACAGATATCC 900
Oy 267 LysSerIleProGlyPheValAsnLeuAsnLeuAsnAspGlnValThrLeuLeuLysTyr 286
Db 901 AAAACCATTCCTGGTGTGTAAATCTGACTGAACGACCAAGTAACCTCTCTCAATATAT 960
Oy 287 GlyValHisGlnIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeu 306
Db 961 GGAGTCCACGAGATCATTTACACAATGCTGGCTCTTGATGAATAAAGATGGGTTTCTC 1020
Oy 307 IleSerGlnGlnGlnGlyPheMetThrArgGlnPheLeuLysSerLeuArgLysProPhe 326
Db 1021 ATATCCGAGGCGCCAGAGCTTCATGCAAGGAGTTCCTAAAGAGCTCGGAAAGCTTTT 1080
Oy 327 GlyAspPheMetGlnProLysPheGlnPheAlaValLysPheAsnAlaLeuGlnLeuAsp 346
Db 1081 GGTGACTTATGAGACCCCAAGTTTGTGAGTTGCTGTGAAGTTCAATGCACTGGAATTAGAT 1140
Oy 347 AspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeu 366
Db 1141 GACACCGACTGTGGCAATATTTATGCTGTCTCATTTCTCAGGGAGACCGCCAGGTTTG 1200
Oy 367 LeuAsnValLysProIleGlnAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGln 386
Db 1201 CTGATGTGGAAGCCCATTTGAGAGACTTCAGACAACTGCTACAGAGCCCTGAGAGTCCAG 1260
Oy 387 LeuLysLeuAsnHisProGlnSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThr 406
Db 1261 CTGAAGCTGAACCACTGAGTCTTCACAGGCTGTTCGCAAGCTCTCCAGAAATGTACA 1320
Oy 407 AspLeuArgGlnIleValThrGlnHisValGlnLeuLeuGlnValIleLysLysThrGln 426
Db 1321 GACCTCAGACGATTTGTACGGAACACGTGCAGCTACTGCAAGGTGATCAAGAAGCGAG 1380

```

```

Oy 427 ThrAspMetSerLeuHisProLeuLeuGlnGlnIleTyrLysAspLeuTyr 443
Db 1381 ACAGACATGAGTCTTACACCCGCTCTCGAGAGATCTACAGAGACTGTGTAC 1431
RESULT 11
AAK36522
ID AAK36522 standard; cDNA: 1811 BP.
AC AAK36522;
XX
XX 07-JUL-1999 (first entry)
DT
XX
XX Human PPAR-gamma1 coding sequence.
DE
XX
XX Nuclear receptor agonist; antagonist; identification; PPAR;
KW peroxisome proliferator activated receptor; ss.
XX
XX Homo sapiens.
OS
XX MO9918124-A1.
PN
XX
XX 15-APR-1999.
PD
XX
XX 06-OCT-1998; 98WO-US21049.
PE
XX
XX 07-OCT-1997; 97US-0061385.
PR
XX
XX (MERI ) MERCK & CO INC.
PA
XX
XX Cummings RT, Hermes JD, Moller DE, Zhou G;
PI WPI: 1999-263998/22.
DR
XX
XX P-PSDB; AAY05471.
DR
XX
XX Identifying nuclear receptor agonists and antagonists
PT
XX
XX Disclosure: Fig 9b; 60pp; English.
PS
XX
XX This sequence encodes the human peroxisome proliferator activated
CC receptor-gamma1 (PPAR-gamma1).
CC
CC The invention relates to a method for identifying nuclear receptor
CC agonists and antagonists comprising measuring fluorescent resonance energy
CC transfer between fluorescent-labelled nuclear receptors and
CC co-activators. The method can be used for identifying agonists and
CC antagonist of nuclear receptors.
XX
XX Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 4,95e-182 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2
US-09-931-007a-1 (1-688) x AAK36522 (1-1811)
Oy 23 AlaAsnIleSerGlnGlnMetThrMetValAspThrGlnMetProPheTyrProThrAsn 42
Db 155 GCGCCGCTGGCCGCGCAATGACCATGGTGTGACACAGAGATCGCATTCCTTGTGATATTC 214
Oy 43 PheGlyIleSerSerValAspLeuSerValMetGlnAspHisSerHisSerPheAspIle 62
Db 215 TTGGGATCAGCTCCGTGATCTCTCGTAAATGGAAGCACTCCCACTCTTGTGATATTC 274
Oy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGlnAspIlePro 82
Db 275 AAGCCCTTCACTACTAGTGTGACTTCTCCGCACTTCTACTCATCTAGAGACATTTCCA 334
Oy 83 PheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysLeuGlnGlnTyr 102
Db 335 TTCAACAAGACAGATCCAGTGTGTCAGATTAACAAGTATGACCTGAAACTCAAGAGATAC 394

```



```

Oy 103 GlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGlnLysThrGlnLeu 122
    |||||||
Db 395 CAAAGTGCATCAAAATGAGGAGCTGCATCTCCACCTTATTATCTGAGAAGACATGCTC 454
Oy 123 TyrAsn----- 124
    |||||||
Db 455 TACAAATTAAGCCTCAGTAAGAGCCCTTCCAACTCCCTCATGGCAATTGAATGCTGCTGTGT 514
Oy 124 ----- 124
Db 515 GGAGATAAAGCTTTCGGATTTCACATATGAGATTGATGCTTGGAAGATGCAAGGCTTTC 574
Oy 124 ----- 124
Db 575 TTCGGAGAACATCAGATTGAAGCTTATCTATGACAGATGATGCTTAAGTCTGCGCATC 634
Oy 125 -----ArgAnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
    |||||||
Db 635 CACAAAAAAGTAGAATAATATGTCTAGTACTGCTCGTTTCAGAAATGCCCTTGACAGTGGGG 694
Oy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeu 160
    |||||||
Db 695 ATGTTCTAATAAGCCATCAGGTTTGGCGGATCGCACAGGCGAGAGAGAGAGAGCTGTTG 754
Oy 161 AlaGlnIleSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg--Ala 179
    |||||||
Db 755 GCGAGATCTCCAGTGAATTCGACGACGATCCAGATCCGCTGACCTCGCTGACGGCC 814
Oy 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
    |||||||
Db 815 CTGGCAAAACATTTGTATGACTCATACATAAAGTCTTCCCTGACCAAAACAAAGCGCG 874
Oy 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
    |||||||
Db 875 AGGGCATTTTACAGAGAGAACAGACAGACAAATCCACATTCGTATCTATACATGAAT 934
Oy 220 SerLeuMetMetGlyGlnAspLysIleLysPheLysHisIleThrProLeuGlnGln 239
    |||||||
Db 935 TCCTTAATGATGGAGAACATTAATCAAGTCAAAACATCACCCTCGACGAGGACG 994
Oy 240 SerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlnAlaVal 259
    |||||||
Db 995 ACCAAAGAGGTGGCCATCGCATCTTTCAGGGCTCCCAATTCCTCCCTGAGGCTGTGG 1054
Oy 260 GlnGlnIleThrGlnTyrAlaLysSerIleProGlnPheValAsnLeuAspLeuAsn 279
    |||||||
Db 1055 CAGGAGATCACAGAGATATCCAAAAGCATTCCTGGTTTGTAAATCTTGACTTGAAACGAC 1114
Oy 280 GlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeu 299
    |||||||
Db 1115 CAAGTAACCTCTCCCAATATATGAGTCCACGAGATCATTTACACATGCTGCTCCCTTG 1174
Oy 300 MetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGlnPheLeu 319
    |||||||
Db 1175 AAGAAATAAAGAGGGGTTCTCATATCCGAGGGCCAAAGGCTTATACAAAGGAGAGTTTCTTA 1234
Oy 320 LysSerLeuArgLysProPheGlnAspPheMetGlnProLysPheGlnPheAlaValLys 339
    |||||||
Db 1235 AAGAGCTTCGCAAGACCTTTTGGTACCTTTATGAGACCCCAAGTTTGAAGTTTCTGTGAAG 1294
Oy 340 PheAsnAlaLeuGlnLeuAspSerAspSerAspLeuAlaIlePheIleAlaValIleLeu 359
    |||||||
Db 1295 TTCAATGCACATGCAATTGATGACAGCGACTGGCAATATTAATTCGTCTATTCTTC 1354
Oy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGlnAspAsnLeu 379
    |||||||
Db 1355 AGTGAGAGCCGCCAGGTTTGTCTGAATGGAAGCCCATTTGAAGACATTCAGACCAACCTTG 1414
Oy 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnLysSerGlnLeuPheAla 399
    |||||||
Db 1415 CTACAAGCCCTGGAGCTCAGCTGAAGCTGAACCAACCTCGAAGTCTTCAAGCTGTTTGCC 1474

```

```

Oy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeu 419
    |||||||
Db 1475 AAGCTCTCCACAAAATGACAGACCTCAGACAGATTGTCACGGAACCTGCAGCTACTG 1534
Oy 420 GlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnIleTyr 439
    |||||||
Db 1535 CAGGTGATCAAAAGACGAGACACATGATGCTTTCACCTCTCTGACGAGATCTAC 1594
Oy 440 LysAspLeuTyrAlaIlePheAlaIleLeuThr 449
    |||||||
Db 1595 AAGGACTTGTACTAGACAGAGACTCCTGAGC 1624

```

## RESULT 12

```

AAH76296
ID AAH76296 standard; cDNA; 1811 BP.
XX
AC AAH76296;
XX
XX
DT 29-OCT-2001 (first entry)
XX
DE Human PPARgamma cDNA sequence.
XX
KW PAX8-PPARgamma1; oncogene; cytosolic; PAX8; PPARgamma1; cancer;
XX follicular carcinoma; PPARgamma; human; ss.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 173..1609
    /*tag= a
XX
PN WO200152789-A2.
XX
PD 26-JUL-2001.
XX
PF 18-JAN-2001; 2001WO-US01664.
XX
PR 20-JAN-2000; 2000US-0177109.
XX PR 14-AUG-2000; 2000US-0225079.
XX
PA (BGHM ) BRIGHAM & WOMENS HOSPITAL INC.
XX
PI Kroll TG, Fletcher JA;
XX
DR WPI: 2001-514487/56.
XX
DR P-PSDB: AAB85802.
XX
PT New PAX8-PPARc1 oncogene and oncoprotein, useful for detecting and
PT treating certain tumors or cancers, e.g. follicular carcinoma -
XX
PS Disclosure: Page 137-139; 145pp; English.
XX
CC The invention relates to an oncogene designated PAX8-PPARgamma1 that
CC contains a PAX8 coding region fused to PPARgamma1 coding region. The
CC PAX8-PPARgamma1 polypeptides can be expressed by standard recombinant
CC methodology. A PPARgamma1 ligand or agent is useful for treating a
CC subject having a disorder characterized by the presence of a PAX8-
CC PPARgamma1, where the disorder is cancer, e.g. follicular carcinoma.
CC The PAX8-PPARgamma1 molecules are also useful for providing nucleotide
CC and amino acid sequences useful for detecting the above disease. The
CC present sequence represents a human PPARgamma polypeptide encoding cDNA.
XX
SQ Sequence 1811 BP; 510 A; 433 C; 422 G; 446 T; 0 other;

```

## Alignment Scores:

```

Pred. No.: 4,95e-182 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: 22 Gaps: 2

```

US-09-931-007A-1 (1-688) x AAH76296 (1-1811)











```

Db 637 ATGAATAAAGATGGGTTCTCATATCCGAGGCCAAGGCTTCATGACAAAGGAGTTCTA 578
QY 320 LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys 339
|||
Db 577 AAGAGCTCGCGAAGCCCTTTGGTGACTTATGGAGCCCAAGTTTGAGTTGCTGTGAAG 518
QY 340 PheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
|||
Db 517 TTCATATGCACGTGAATAGATGACACGCACTTGGCAATATTATTGCTGTCATTATTCTC 458
QY 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379
|||
Db 457 AGTGGAGACCGCCCGAGTTGCTGAATGTGAAGCCCATTTGAAGACATTCAGACCAACCTG 398
QY 380 LeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399
|||
Db 397 CTACAGCCCTGGAGCTCCAGCTGAAGCTGAACCACTGAGTCTTCACAGCTGTTGCC 338
QY 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeu 419
|||
Db 337 AAGCTGCTCCAGAAATGACAGACTCAGACAGATTGTCACGGAACAGCTGCAGCTACTG 278
QY 420 GlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyr 439
|||
Db 277 CAGGTGATCAGAGAGCGAGACGACATGAGTCTTCACCCGCTCCTGCAGGAGATCTAC 218
QY 440 LysAspLeuTyrAlaTrpAlaIleLeuThr 449
|||
Db 217 AAGGACTTGTAAGAGAGAGAGTCTGAGC 188

```

Search completed: February 25, 2003, 04:09:13  
 Job time : 389 secs



GenCore version 5.1.3  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame\_plus.p2n model

Run on: February 25, 2003, 04:02:01 ; Search time 2471 Seconds

(Without alignments)  
4509.306 Million cell updates/sec

Title: US-09-931-007A-1  
Perfect score: 3508  
Sequence: 1 MGFELGSDPIDESDSFTDR.....KTEIDMSLHDLQIYKLDY 688

Scoring table:  
BLOSUM62  
Xgapop 10.0, Ygapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n model -DPY=xjh  
-Q=/cgn2\_1/USPFO/US09931007/runat\_14022003\_102553\_4687/app.query.fasta.1.839  
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45  
-DOCLALIGN=200 -THR\_SCORE=pcct -THR\_MAX=100 -THR\_MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09931007 @CGN.1.1698 @runat\_14022003\_102553\_4687 -NCPU=6 -ICPU=3  
-NO\_XLPPY -NO\_MMAP -LARGEOUFRY -NEB\_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120  
-MARK.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -YGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7  
-YGAPOP=10 -YGAPEXT=0.5 -DELPEXT=7

Database :  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estlin:\*  
4: em\_estnu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_htc:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_htc:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_esthum:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_iny:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_rtd:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1341	38.2	961	9	AL549707
2	1318	37.6	843	9	AL570116
3	1258	35.9	913	9	AL573383
4	1253	35.7	932	9	AL571005
5	1199	34.2	1026	9	AL523433
6	1198	34.2	793	12	BG742110
7	1136	32.4	826	10	BE535401
8	1122.5	32.0	900	9	AL549611
9	1098.5	31.3	1088	12	BG179310
10	1060.5	30.2	985	9	AL545177
11	996.5	28.4	944	14	BO883138
12	985	28.1	903	9	AL543579
13	985	28.1	872	14	BQ722874
14	933.5	26.6	901	13	BI820841
15	924.5	26.4	898	9	AL549037
16	919.5	26.2	574	9	AA625223
17	909.5	25.9	865	12	BG534958
18	905	25.8	551	12	BF192896
19	865	24.7	911	14	BO930349
20	846	24.1	513	12	BF192893
21	837.5	23.9	959	9	AL523434
22	833	23.7	888	12	BG166774
23	813	23.2	719	14	BQ781472
24	812.5	23.2	932	14	BO943112
25	798	22.7	694	14	BQ021210
26	767.5	21.9	1011	12	BE904444
27	757.5	21.6	986	14	BM923992
28	753.5	21.5	592	10	AM886988
29	752.5	21.5	700	9	AA673643
30	751.5	21.4	734	14	BO571382
31	740	21.1	450	13	BM433468
32	722	20.6	478	9	AI037566
33	716	20.4	533	10	AV605730
34	715	20.4	815	10	BE300340
35	699.5	19.9	938	10	BE300333
36	687	19.6	432	9	AA717831
37	684	19.5	746	9	AU131142
38	683	19.5	654	14	BM995901
39	682	19.4	965	14	BQ947830
40	681	19.4	430	12	BF522355
41	653	18.6	659	13	BI066593
42	649	18.5	823	13	BI524663
43	647	18.4	749	13	BI186429
44	646	18.4	524	10	AW911965
45	641.5	18.3	888	12	BG436329

#### ALIGNMENTS

RESULT 1  
AL549707/c  
LOCUS AL549707 LTL.NFL006.Pl2 Homo sapiens cDNA clone CS001055Y11.3  
DEFINITION prime, mRNA sequence.  
ACCESSION AL549707  
VERSION AL549707.1 GI:12865958  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 961)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization

JOURNAL Unpublished (2001)  
COMMENT Contact: Genoscope

COMMENT

Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91106 EVRY cedex - France  
Email: [segref@genoscope.cns.fr](mailto:segref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

## FEATURES

```
source      1. .961
```

```

1. .961
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD105Y111"
/clone_lib="LTI_NF1006_PL2"
/rnause_type="placenta"
/notes=Vector: PCWMSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pcwmsport 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifestech.com URL :
http://fulllength.invitrogen.com"

```

[illegible]

ORIGIN

Alignment Scores:	
Pred. No.:	3,56e-160
Score:	1341.00
Percent Similarity:	95.49%
Best local Similarity:	94.10%
Query Match:	38.23%
DB:	9
US-09-931-007A-1 (1-688) x AL549707 (1-961)	
	Length: 961
	Matches: 271
	Conservative: 4
	Mismatches: 12
	Indels: 0
	Gaps: 0

US-09-931-007A-1 (1-688) x AL549707 (1-961)

QY	163	IIleSerAspIleAspGlnIleuAsnProGluSerIleAspLeuArgAlaIleuAlaIys	182
Db	953	GTTCACAGTGAATACGACCAAGCTAAATCCAGATGCCGTACCTCCGGCCCTGGCAAAA	894
QY	183	HisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaIysAlaArgAlaIle	2023
Db	893	CATTGTATACATCATACATAAAATGCTCTCCCGCTGCACAAAGMAAGGAGGGGATC	834
QY	203	LeuThrGlyLysThr-ThrAspLysSerProPheValIleTyrAspMetLsnSerLeuWe	222
Db	833	TTGCACAGAAAGACAMACAGCAAAATACACATTCGGTATTATAGACTGTAATTCCTTAAT	774
QY	222	LMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGlnGluInsSerIysG	242
Db	773	GATGGAGAAAGATTAATCAAGTTCAACAAACATCACCCCTCCAGAGACGACAAAG	714
QY	242	uValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaGlnGluI	262
Db	713	GGTGGCCATCCGGATCTTTTCAGGGCTGCCGGTTCGGTCCGTGGAGGCTGTGCAGAGAT	654
QY	262	eThrGlyTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValTh	282
Db	653	CACGAGATATGCCAAAGCATTCCTGCTTTTGTAAATCTTGAACTTGACCTTGACCAACCAAGTAA	594
QY	282	rLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetLsnIy	302
Db	593	TCTCTCAAAATATGAGAGTCCACGAGATCATTTACACAACTCTGGCTCTGTATGAATTA	534
QY	302	sAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSerIe	322
Db	533	AGATGGGGTTCATATACCGAGGGCCCAAGGCTTCATGACACAGGAGATTTCTTAAGAGCCT	474
QY	322	uArgLysProPheGlyLysPheMetGluProLysPheGlnPheAlaValLysPheAsnI	342
Db	473	GCGAAACCTTTTGGTACTTTATATGAGGCCCAAGTTGAGTTTGGCTGTGAAGTTCAATGC	414
QY	342	AlaGluLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAs	362

Db	413	ACTGGAAATTAGATGACAGCGCACTGGCGAATATTATTATGGTGTGATTATTCTCACTGAGTGGAGA	354
OY	362	PATGProGlyLeuLeuLeuSnaValLysProIleGluAspIleGlnAspSnaLeuLeuGlnAl	382
Db	353	CGGCCACAGTTTGGTGTGAATGTGAAGCCCATTTGAAGRCATTTCAAGACACACCTGGCTTACAAAC	294
OY	382	aleuGluLeuGlnLeuLeuLysLeuSnaHisProGlnLysSerGlnLeuPheAlaLysLeu	402
Db	293	CCTGGAGCTCCMGCTGAAAGCTGAAACCAACCTGAGATCCTCAAGCTGTTTGCCAAAGCTGCT	234
OY	402	uGlnLysMetThrAspLeuAlaTgGlnIleValThrGlnHisValGlnLeuLeuGlnValI	422
Db	233	CCAGAAATATKACMGACCTCAACACGATGTTCMGCAACMGCTGCGTACTGACAGGTGAT	174
OY	422	eLysLysThrGlnLuthrAspMetSerLeuHisProLeuLeuGlnGluLeuTyrLysAspLe	442
Db	173	CAAGAAAGCGAGACAGACATGATCTTCAACCCGCTCTCGACAGAGATTCACAAAGACTT	114
OY	442	uTyrAlaTrpAlaIleLeuThr	449
Db	113	GTACTGWCAGAGAGTCTGTGAGC	92

[illegible]

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1. (bases 1 to 843)	Li, W. B., Gruber, C., Jesse, J., and Polares, D.	Full-length cDNA libraries and normalization	Unpublished (2001)	Contact: Genoscope Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES	Location/Qualifiers
source	1. .843

/organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CSDD10067110"  
 /clone\_1b="LTI\_NF1006\_PL2"  
 /tissue\_type="placenta"  
 /note="Vector: pCMVSPORT6; Site 1: NotI; first strand cDNA was primed with a NotI-clo/drf(1) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : <http://fulllength.invitrogen.com>"

BASE COUNT	195 a	196 c	208 g	239 t	5 others
ORIGIN					

## ORIGIN

Alignment Scores:	
Pred. No.:	2.48e-157
Score:	1318.00
Percent Similarity:	97.44%
Best Local Similarity:	96.70%
Query Match:	37.57%
DB:	9
Length:	84
Matches:	26
Conservative:	2
Mismatches:	7
Indels:	0
Gaps:	0

US-09-931-007A-1 (1-688) x AL570116 (1-843)

177 LeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLys 196





```
|||||
Db 312 CCACCCGAGCTCCTCAGAGCTGTTGGCCAGCGCTCCACAAAATGACACACCTCAGACA 253
OY 410 nllvalThrgLHHisValGlnleuenglnvalilleyLysThrGlnThrAspMetse 430
Db 252 GATTTCACGGAACACCTGAGGACTGACGAGTGATCAAGAGACGAGACATGAG 193
OY 430 lIeuHisProleuLeuGlnluileTyLysAspLeuTyAlaTrpAlaileuThr 449
Db 192 TCTTCAACCCCTCCTGAGAGATCTCAAGACTTCTACTACAGAGACTCCTGAGC 135

RESULT 4
AL571005/ 932 bp mRNA linear EST 16-FEB-2001
LOCUS AL571005 LTI_NFL006.PL2 Homo sapiens cDNA clone CS0D1028YJ24 3
DEFINITION prime, mRNA sequence.
ACCESSION AL571005
VERSION AL571005.1 GI:12927873
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 932)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT Genoscope - Centre National de Sequencage
Genoscope BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES
source 1..932
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL006.PL2"
/tissue_type="placenta"
/feature="Vector: PCWSPORT 6; Site 1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with NotI and
cloned into the NotI and EcoRV sites of the PCWSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 230 a 207 c 210 g 281 t 4 others
ORIGIN
Alignment Scores:
Pred. No.: 6e-149 Length: 932
Score: 1253.00 Matches: 233
Percent Similarity: 97.70% Conservatave: 2
Best Local Similarity: 96.93% Mismatches: 5
Query Match: 35.72% Indels: 1
DB: 9 Gaps: 0
US-09-931-007a-1 (1-688) x AL571005 (1-932)
OY 190 LysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrAsp 209
Db 919 AAKCCTCCCGCTGACCAAGCAAGGCGGCGATCTTGACAGGAAAGACACAGAC 860
OY 210 LysSerProPheValIleTyAspMetAsnSerLeuMetLysLysLysLysLys 229
Db 859 AAATCCCAATCTCTATATGATGACGAATTCCTTAATGATGGAGAGATAAATCAAG 800
OY 230 PheLysHisIleThrProLeuGlnGlnGlnIleSerLysGluValAlaIleArgIlePheGln 249
Db 799 TTCAAACACATCCACCCCTGCGAGAGACAGCAAGAGGTGGCATCCGCATCTTTGAG 740
```

```
OY 250 GlyCysGlnPheArgSerValGluAlaValGlnluileThrGluTyAlaLysSerIle 269
Db 739 GGCTGCACAGTTTCGCTCCGCGAGAGGCTGTGCAGAGATCATAGATATGCGCAAAACAT 680
OY 270 ProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyArgValHis 289
Db 679 CCTGGTTTGTAAATCTTGTGATGAACGACCAAGTAACTCTCTCAAAATGAGATCCAC 620
OY 290 GluIleIleTyThrMetLeuAlaSerLeuMetAsnLysAspGlyValIleIleSerGln 309
Db 619 GAGATCATTTTACACAATGCTGGCTCTCTTGATCAATTAAGATGGGTTCTCATATCCGAG 560
OY 310 GlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAspPhe 329
Db 559 GGCCAGAGGCTCATGACAAGGAGGTTTCTTAAGAGCCTGGAAAGCCTTTTGGTGACTTT 500
OY 330 MetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGlnLeuAspAspSerAsp 349
Db 499 ATGAGACCCCAAGTTTGAAGTTTGTGAGTTCAATGACACTGGAATTAATGATGACACCGAC 440
OY 350 LeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnVal 369
Db 439 TTGGCAATATTTTGTCTGTCTATATTCTCAGTGGAGACCGCCAGCTTGCTGAATGTG 380
OY 370 LysProIleGluAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeu 389
Db 379 AAGCCCATTTGAACATTCACAGCAACCTGCTCAAGCCCTGGAGCTCCAGCTGAAGCTG 320
OY 390 AsnHisProGluSerSerGlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArg 409
Db 319 AACCAACCTGAGCTCTTACAGCTGTTGGCCACCTCTCCAGAAATGACAGACCTCAACA 260
OY 410 nllval-ThrgLHHisValGlnleuenglnvalilleyLysThrGlnThrAspMetse 429
Db 259 CAGATTGTACAGGACGACACGCTGACGACTGACGAGTGATCAGAGACGAGACGACAT 200
OY 429 tSerLeuHisProleuLeuGlnluileTyLysAspLeuTyAlaTrpAlaileuThr 449
Db 199 GAGTCTTCAACCCCTCCTGAGAGATCTCAAGACTGTACTACGAGAGGTCTGAG 140
OY 449 t 449
Db 139 c 139

RESULT 5
AL523433/ 1026 bp mRNA linear EST 13-FEB-2001
LOCUS AL523433 LTI_NFL003_NBC3 Homo sapiens cDNA clone CS0DC004Y008 3
DEFINITION prime, mRNA sequence.
ACCESSION AL523433
VERSION AL523433.1 GI:12786926
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1026)
Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
COMMENT Genoscope - Centre National de Sequencage
Genoscope BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES
source 1..1026
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="LTI_NFL003_NBC3"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
```



QY 204 ThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsnSerLeuMetMet 223  
|||||  
Db 277 ACAGGAAGAAACAAACAGACAAATCACCATTGCTATCTATGACATGCAATTCCTTAATGATG 336  
QY 224 GlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGlnGlnInseryGluVal 243  
|||||  
Db 337 GGAGGAAGATAAATCAAGTTCAACACATCACCCCTCAGAGAGAGCAAGAGAGTG 396  
QY 244 AlaIleArgIlePheGlnLysCysGlnPheArgSerValGluValAlaGlnGluIleThr 263  
Db 397 GCCATCCGATCTTTCAGGGCTGCCAGTTCCGCTCGTGAGAGCTTGACAGAGATCACA 456  
QY 264 GluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeu 283  
|||||  
Db 457 GAGTAAAGCCAAAGCATCTGCTGTTTGAATCTTGACTGAAAGCAGCAAGAACTCTC 516  
QY 284 LeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAsp 303  
|||||  
Db 517 CTCGAATATGAGAGTCCAGACATCATTACACATCTGCTGCTGCTGATGATTAAGAT 576  
QY 304 GlyValLeuIleSerGluGlnGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArg 323  
Db 577 GGGGTTCTCATATCCAGAGGCCAAGGCTTCATGACAGAGAGTTCTTAAGAGCCTCGCA 636  
QY 324 LysProPheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeu 343  
Db 637 AAGCCTTTGGTGACTTTTATGAGCCCAAGTTGAGCTTGCTGTGAAGTTCAATGCACATG 696  
QY 344 GluLeuAspAspSerPheLeuAlaIle-PheIleAla-ValIleIleLeuSerGlyAspA 363  
Db 697 GAATTAAGATGACAGCAGCTTGGCAATATTTTATGCTGCTCATATTCTCAGTGAGACC 756  
QY 363 rgProGlyLeuLeuAsnValLysProIleGluAspIle 375  
Db 757 G-CCAGGTTTC-CTGAATGTGAAGCNCATTTGAAGACATT 792  
RESULT 7  
LOCUS BE535401 826 bp mRNA linear EST 09-AUG-2000  
DEFINITION 60105854F1 NIH\_MGC\_10 Homo sapiens cDNA clone IMAGE:3445380 5',  
mRNA sequence.  
ACCESSION BE535401  
VERSION BE535401.1 GI:9764046  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 826)  
NIH-MGC http://mgc.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: Incyte Genomics, Inc.  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNI at:  
http://image.llnl.gov  
Plate: LAM8415 row: k column: 13  
High quality sequence stop: 736.  
Location/Qualifiers  
1. 826  
/organism="Homo sapiens"  
/db\_xref="taxon:9606"  
/clone="IMAGE:3445380"  
/clone\_lib="NIH\_MGC\_10"  
/cell\_line="MGC36"  
/lab\_host="DH10B"  
/note="Organ: cervix; Vector: pCMV-Sport6; Site\_1: NciI;  
Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dT.

BASE COUNT 229 a 187 c 221 g 189 t  
ORIGIN  
Alignment Scores:  
Pred. No.: 4,42e-134 Length: 826  
Score: 1136.00 Matches: 248  
Percent Similarity: 94.68% Conservative: 1  
Best Local Similarity: 94.30% Mismatches: 11  
Query Match: 32.38% Indels: 7  
DB: 10 Gaps: 0  
US-09-931-007A-1 (1-688) x BE535401 (1-826)  
QY 129 GlnTyrCysArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArgPhe 148  
Db 2 CAGTACTGTCGGTGCAGAAATGCTTCGACATGGGATCTCATATATCCATCAGTTT 61  
QY 149 GlyArgMetProGlnAlaGluLysGluLysLeuLeuAlaGluIleSerSerAspIleAsp 168  
Db 62 GGGCGGATGCCACAGGCCAGAAAGAGAGCTGTGGCGGAGATCTCCAGTATATCGAC 121  
QY 169 GlnLeuAsnProGluSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyr 188  
Db 122 CAGCTGAATCCAGAGTCCCTGACCTCCGCGCCCTGGCAAAACATTGTATGACTCATAC 181  
QY 189 IleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlnLysThrThr 208  
Db 182 ATTAAGTCTTCCCTCCGCTGACCAAGCAAGAGGCGCATCTTGACAGAAAGACACACA 241  
QY 209 AspLysSerProPheValIleTyrAspMetAsnSerLeuMetMetGlyGluAspLysIle 228  
Db 242 GACAAATCACCATTGCTGTATGATGACATGAATTCCTTATGATGGAGAAATAAATC 301  
QY 229 LysPheLysHisIleThrProLeuGlnGlnInseryGluValAlaIleArgIlePhe 248  
Db 302 AAGTTCAAAACATCACACCTCCCTGACAGAGACAGACAAAGAGTGCCATCCGATCTTT 361  
QY 249 GlnGlyCysGlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAlaLysSer 268  
Db 362 CAGGGCTGCCAGTTTGCCTCCCTGAGGCTGTGCAGAGATCACAGATATCCAAAGCC 421  
QY 269 IleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyVal 288  
Db 422 ATTCTGCG-TTTGTAATCTTGACTGAACGACCAAGTAACTCTCTCAATATGAGATC 480  
QY 289 HisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSer 308  
Db 481 CACGAGATCATTTACACATGCTGGCTCTTGATGAATAAAGATGGGTTCTCATATCC 540  
QY 309 GluGlyGln-GlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPhe-GlyA 328  
Db 541 GAGGGCCAGAGGCTTCATGACAAAGGAGTTCTTAAGAGCCGTGCAGAAACCTTTGGG 600  
QY 328 spPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGlnLeuAspAsp 348  
Db 601 ACTTATGAGGCCAAAGTTGAGTTGGCTGCTGATGATCATGACATGGAATTAAGTAC 660  
QY 348 erAspLeuAlaIlePheIle-AlaValIleIleLeuSerGlyAspArgProGlyLeuLeu 367  
Db 661 GGGACTTGCAATATTTATGCGGGCCATTATCTCAGGGAGAGCCG-CCAGCTTGGCTG 719  
QY 368 AsnValLysProIleGlnAspIleGlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeu 387  
Db 720 AATGTGAAGCCATT-GAAGCATTCAGGACAAACATGCTTCAAGCCG-GAGGCTACGCTG 777  
QY 388 Lys 388  
Db 778 AAG 780  
RESULT 8  
AL549611

LOCUS AL549611 900 bp mRNA linear EST 16-FEB-2001  
 DEFINITION AL549611 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D105Y111 5  
 prime, mRNA sequence.  
 ACCESSION AL549611  
 VERSION AL549611.1 GI:12885763  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 900)  
 AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.  
 FEATURES  
 source  
 1. .900  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="CS0D105Y111"  
 /tissue\_type="placenta"  
 /note="Vector: pCMVSPORT 6; site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Lite Technologies. Contact : Feng Liang Life Technologies, a division of invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 266 a 201 c 214 g 218 t 1 others  
 ORIGIN

Alignment Scores:  
 Pred. No.: 2.72e-132 Length: 900  
 Score: 1122.50 Matches: 229  
 Percent Similarity: 77.78% Conservative: 2  
 Best Local Similarity: 77.10% Mismatches: 3  
 Query Match: 32.00% Indels: 63  
 Gaps: 2

US-09-931-007A-1 (1-688) x AL549611 (1-900)

OY 94 LysTYrAspLeuLysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSerPro 113  
 |||||  
 Db 3 AAGTATGACCTGAAACCTCAAGAGTACCAAGTCAATCAAGTGGAGCTGCACTCTCCA 62  
 |||||  
 OY 114 ProTYrSerGluLysThrGlnLeuTyrAsn----- 124  
 |||||  
 Db 63 CCTATTATTCTGAGAAAGCTCAGCTCAATATAGCCTCATGAAGAGCTTCCAACTCC 122  
 |||||  
 OY 124 ----- 124  
 Db 123 CTCATGCAATTGAATGTCGTCTGTGGAGATAAAGCTTCTGATTCATGAGACTT 182  
 |||||  
 OY 124 ----- 124  
 Db 183 CATGCTGTGAAGATGCAAGGTTTCTCCGAGAACATCAGATTGAACCTTATCTAT 242  
 |||||  
 OY 125 -----ArganLysCysGlnTyrCys 131  
 |||||  
 Db 243 GACAGATGTGATCTTAAGTGTGATCCACAAAAAAGTAAAGTAATGTCAGACTGT 302  
 |||||  
 OY 132 ArgPheGlnLysCysLeuAlaValGlyMetSerHisAsnAlaIleArgPheGlyArgMet 151  
 |||||  
 Db 303 CGGTTTCAGAAATGCGCTTGCAGTGGGATGTCTCATATATGCCATCAGGTTTGGCGGATG 362  
 |||||  
 OY 152 ProGlnAlaGluLysGluLysLeuAlaGluIleSerSerAspIleAspGlnLeuAsn 171

Db 363 CCACAGGCCGACGAGACGAGAACTCTTGCGGAGATCTCCAGTATACGACACCTGAAT 422  
 |||||  
 OY 172 ProGlnSerAlaAspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSer 191  
 |||||  
 Db 423 CCAGATGCGGTGACCTCCGGCCCTGGCAAAACATTTGTATGACTCATATCAATCAATCC 482  
 |||||  
 OY 192 PheProLeuTyrLysAlaLysAlaArgAlaIleLeuTyrGlyLysThrThAspLysSer 211  
 |||||  
 Db 483 TTCCCGCTGACCAAGCAAGCCAGCGCATCTTGACACGAAAGACACACGACAAATCA 542  
 |||||  
 OY 212 ProbheValIleTyrAspMetAsnSerLeuMetGlyGluAspLysIleLysPheLys 231  
 |||||  
 Db 543 CCATTGCTTATCTATGACATGAATCTTATATGATGAGAGAAAGATTAATCAATCTCAA 602  
 |||||  
 OY 232 HistLeuThrProLeuGlnGlnSerLysLysLysValAlaIleArgIlePheGlnGlyCys 251  
 |||||  
 Db 603 CACATCACCCCTGACAGACAGACAAAGAGTGGCCATCCCATCTTTCACGCGCTGC 662  
 |||||  
 OY 252 GlnPheArgSerValGluAlaValGlnGluIleThrGluTyrAlaLysSerIleProGly 271  
 |||||  
 Db 663 CAGTTTCCCTCCGTGAGAGCTGTGACAGATCACAGATGTGCCAAAGCATTCCTGTGT 722  
 |||||  
 OY 272 PheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIle 291  
 |||||  
 Db 723 TTTGTAATCTTGACTTGAAAGACCAAGTACTCTCCCAATATGAGATCCACAGATC 782  
 |||||  
 OY 292 IleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlnGln 311  
 |||||  
 Db 783 ATTTACACAAATGCTGGCTCTTATGATTAAGATGAGGGTTCATATTCAGAGGCCCA 842  
 |||||  
 OY 312 GlyPheMetThrArgGluPheLeuLysSerLeuArgLysProPheGlyAsp 328  
 |||||  
 Db 843 GGCCTTCATGACAAAGGAGGACTTCTAAAGACT---GCGAAASCTTTTGGTGAC 890  
 |||||

RESULT 9  
 BG179310 1088 bp mRNA linear EST 06-FEB-2001  
 LOCUS 602331030F1 NIH\_MGC\_91 Homo sapiens cDNA clone IMAGE:4432278 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BG179310  
 VERSION BG179310.1 GI:12686013  
 KEYWORDS EST.  
 SOURCE human.  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1088)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: rgabs@remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: L14M10188 row: 1 column: 07  
 High quality sequence stop: 658.  
 Location/Qualifiers  
 1. 1088  
 /organism="Homo sapiens"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:4432278"  
 /tissue\_type="adipocarcinoma, cell line"  
 /note="Origin: prostate; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.4 kb. Library enriched for

full-length clones and constructed by life technologies.

Note: this is a NIH\_MGC Library."

BASE COUNT 296 a 255 c 286 g 250 t 1 others

ORIGIN

Alignment Scores:

Pred. No.: 4,41e-129 Length: 1088  
Score: 1098.50 Matches: 242  
Percent Similarity: 96.47% Conservative: 4  
Best Local Similarity: 94.90% Mismatches: 8  
Query Match: 31.31% Indels: 6  
DB: 12 Gaps: 0

US-09-931-007A-1 (1-688) x BG179310 (1-1088)

OY 143 HlSAsnAlaIlleArpPheGlyArqMetProGlnAlaGluLysGluLysLeuAlaGlu 162  
DB 3 CArNAATGCCATACGGTTTGGGGGATGCCAGGCCGAGAGAGAGAGAGCTGTTGGCGGAG 62  
OY 163 lIeSerSerAspIlleAspGlnLeuAsnProGluSerAlaAspLeuArGaLalaLeuAlaLys 182  
DB 63 ArCTCCAGATATGCACAGCGAGATCCAGATCCGCTGACCTCCGGGCCCTGGCAAAA 122  
OY 183 HlSLeuTYrAspSerTYrIlleLysSerPheProLeuThrLysAlaLysAlaArGaLalIle 202  
DB 123 CATTTGTATGACTCATACATCAATAAGTCTTCCGCTGACCAAAAGCGAGAGCGGATC 182  
OY 203 LeuThrGlyLysThrThrAspLysSerProPheValIlleTYrAspMetAsnSerLeuMet 222  
DB 183 TTGACAGGAAAGACACACACAAATCACCATTCGTATATGACATGAAATTCCTTAATG 242  
OY 223 MetGlyLAspLysIlleLysPheLysHlSIlleThrProLeuGlnGluInSerLysGlu 242  
DB 243 ATGGGAGAGATAAATCAAGTTCAACACATCACCCCTCCAGAGACGACGACAAAGAG 302  
OY 243 ValAlaIlleArGIllePheGlnGlnLysGlnPheArGSerValGluAlaValGlnGluIle 262  
DB 303 GTGGCCATCCGATCTTTTCAGGGCTCCAGTTCCGCTCCGTGGAGGCTGCGACGAGATC 362  
OY 263 ThrGluTYrAlaLysSerIlleProGlyPheValAsnLeuAspLeuAsnAspGlnValThr 282  
DB 363 ACAGACTATGCCAAAGACATCCCTGTTTGTAATCTTGACTGAAACACCAAGAACT 422  
OY 283 LeuLeuLysTYrGlyValHlSgluIlleIleTYrThrMetLeuAlaSerLeuMetAsnLys 302  
DB 423 CArCTCAATATATGAGATGCCAGCATCATTTACAAATGCTGGCCCTTGATGATATAA 482  
OY 303 AspGlyValLeuIlleSerGluGlnGlnLysPheMetThrArgGluPheLeuLysSerLeu 322  
DB 483 GATGGGGTCTCATATCCAGGGCCCAAGGCTTCATGACAAAGGAGTTCTTAAAGACCTG 542  
OY 323 ArGlySerProPheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAla 342  
DB 543 CCAAGACCTTTGGTGACTTTATGAGGCCCAAGTTGAGTATNGCTGTGAAGTTCAATGTA 602  
OY 343 LeuGluLeuAspAspSerAspLeuAlaIllePheIlleAlaValIlleLeuSerGlyAs 362  
DB 603 CTGGAAATATGATGACAG-GACTTGGCAATATTTATATGCTGTCATATATCTCCAGTGAGA 661  
OY 362 PArGProGlyLeuLeuAsnValLysProIlleGluAspIlleGlnAspAsnLeuLeuGlnAl 382  
DB 662 CCGGGA-GGGTTGCTGAAGGTGAAGCCATTGAGACTT-CAAGACAACTGTT-AAAGC 718  
OY 382 AlauGluLeuGlnLeuLysLeuAsnHlSProLysSerSerGln 396  
DB 719 TCTGAGAGTCCAGTTGAACTGAACA--CCTGAGTCTCTACAG 759

RESULT 10 985 bp mRNA linear EST 16-FEB-2001  
AL545177  
LOCUS AL545177 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D10281J24 5  
DEFINITION AL545177 LTI\_NFL006.PL2 Homo sapiens cDNA clone CS0D10281J24 5  
ACCESSION AL545177  
prime, mRNA sequence.

VERSION AL545177.1 GI:12877658  
KEYWORDS EST.  
SOURCE human.  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 985)  
AUTHORS Li, W.B., Gruber, C., Jesssee, J., and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)

COMMENT Contact: Genoscope  
Genoscope - Centre National de Sequencage  
BP 191 91006 Evry cedex - France  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 985  
Location/Qualifiers

/organism="Homo sapiens"

/db\_xref="taxon:9606"

/clone="CS0D1028YJ24"

/clone\_1lb="LTI\_NFL006.PL2"

/tissue\_type="Placenta"

/note="Vector: pCMVSPORT 6; Site\_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Peng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : filiang@life.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 290 a 236 c 210 g 247 t 2 others

ORIGIN

Alignment Scores:

Pred. No.: 2.68e-124 Length: 985  
Score: 1060.50 Matches: 229  
Percent Similarity: 74.43% Conservative: 1  
Best Local Similarity: 74.11% Mismatches: 6  
Query Match: 30.23% Indels: 77  
DB: 9 Gaps: 2

US-09-931-007A-1 (1-688) x AL545177 (1-985)

OY 22 SerAlaAsnIlleSerGln-----GluMetThr 30  
DB :::: ||||| ||| |||||  
OY 59 ACCACAAATATACCAAGGCCATTTCGCAACAGACATMAAGCTTTACGAATGACC 118  
DB 31 MetValAspThrGluMetProPheThrProThrAsnPheGlyIlleSerSerValAspLeu 50  
DB 119 ATGTTTGACACAGAGATGCCATTCTGGCCACCACTTTGGGATCAGCTCCGTGATCTC 178  
OY 51 SerValMetGluAspHlSerHlSerPheAspIlleLysProPheThrThValAspPhe 70  
DB 179 TCCGTAATGAGAGACCACCTCCACTCTTTGATATCAAGCCCTTACATGCTGACTTC 238  
OY 71 SerSerIlleSerThrProHlSTyrGluAspIlleProPheThrArgTYrAspProValAl 90  
DB 239 TCCACATTCTTACTCCACATTTACGACAGACATTCATTCCACAGAACAGATCCAGTGT 298  
OY 91 AlaAspTYrLysTYrAspLeuLysLeuGlnGluTYrGlnSerAlaIlleLysValGluPro 110  
DB 299 GCAGATTACAAAGTATGACCTGAACTTCAAGAGTACCAAGTGCAATCAAAAGTGGAGCT 358  
OY 111 AlaSerProProTYrTYrSerGluLysThrGlnLeuTYrAsn----- 124  
DB 359 GCATCTCACCTTATATTTGAGAAGACTCAGCTCTACAAATTAAGCTTCATGAAGAGCT 418  
OY 124 ----- 124  
DB 419 TCCAACTCCCTCATGCAATTGAATGCTGCTGTGAGATAAAGCTTTCGATTTTCAC 478  
OY 124 ----- 124

```
Db 479 TATGAGTTCATGCTTGTGAGAGATGCAAGGGTTCTTCGAGAGAACATTCAGATTGAAG 538
QY 125 -----ArgAsnlyScys 128
Db 539 CTTATCTATGACAGATGATCTTAACCTGCGATCCACAAAAAGTGAATTAATGT 598
QY 129 GlnTyrcysArgphcGlnlyscysLeuAlaValGlyMetSerHisAsnAlaIleArgphe 148
Db 599 CAGTACTGTGGTTTCAGAAATGCCCTTGCAAGTGGGATGTCATATATCCATCAGGTTT 658
QY 149 GlyArgMetProGlnAlaGlnlysglyLeuLeuAlaGlnIleSerSeraspIleasp 168
Db 659 GGGCGGATGCCACAGGCCAGAGAGAGAGCTGTGGCGAGATCTCCAGTATATCGAC 718
QY 169 GlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaValHisIleuTyrcysPhe 188
Db 719 CAGCTGATATCCAGAGTCCGCTTACCTCCGGGCCCTGGCAAAACATTTGTATGACTCATAC 778
QY 189 IleTyrcysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThr 208
Db 779 ATAAAGTCTCTCCGCTGACCAAGCAAGGCGGCGATCTTGACAGAAACAAACA 838
QY 209 AspLysSerProPheValIleTyrcysMetAsnSerLeuMetGlyGlnAspLysIle 228
Db 839 GACAAATCACCATTCGTATCTATGACAGAAATTCCTTAATGATGGAGAAATTAATC- 897
QY 229 LysPheLysHisIleThrProLeuGlnGlnSerLysGlnValAlaIleArgIlePhe 248
Db 898 AAGTTCAACACATCACCCCTG-CAGGAGCAGAGCAAGAGAGTGCCCAT-CCCATCTTT 955
QY 249 GlnGlyCysGlnPheArgSerValGlu 257
Db 956 CAGGCTGCAG-TTTCGCTCCGTGAG 981
RESULT 11
B0883138 944 bp mRNA linear EST 16-AUG-2002
LOCUS AGENCOURT_7982264 lupski_dorsal_root_ganglion Homo sapiens cDNA
DEFINITION B0883138
ACCESSION B0883138
VERSION B0883138.1 GI:22275146
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 944)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health/ Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: c9apbs-rf@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LMNL at:
http://image.llnl.gov
Place: LLM13578 row: 3 column: 11
High quality sequence stop: 707.
Location/Qualifiers
1..944
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6186346"
/clone_id="lupski_dorsal_root_ganglion"
/sex="male"
/tissue_type="dorsal root ganglia"
/dev_stage="adult, 36 yr."
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1:"
```

```
BASE COUNT 282 a 231 c 196 g 234 t 1 others
ORIGIN
NotI: Site-2: SalI: cDNA made by oligo-dT priming.
Directionally cloned using the following adaptors:
5'-TCGACCCACGGCTCG-3' and
5'-GACTATGCTGTGATGCGAGCGCGCCCTT(15)-3', size selected >
1 kb for average insert length 1.7 kb. This is a primary
library, non-amplified. Library constructed by Life
Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
College of Medicine) and is available through life
Technologies."
Alignment Scores:
Pred. No.: 3,78e-116 Length: 944
Score: 996.50 Matches: 204
Percent Similarity: 72.13% Conservative: 3
Best Local Similarity: 71.08% Mismatches: 74
Query Match: 28.41% Indels: 74
DB: 14 Gaps: 2
US-09-931-007a-1 (1-688) x B0883138 (1-944)
QY 22 SerAlaAsnIleSerGln-----GlnMetThr 30
Db 85 ACCACAAATATATCAACAAGGCCATTTCCTCAACGAGATGACCTTTAAGCAAAATGACC 144
QY 31 MetValAspThrGlnMetProPheThrProHisPheGlyIleSerSerValAspLeu 50
Db 145 ATGCTTGACACAGATATGCAATTCCTGGCCACCACTTTGGGATGACCTCCGTGATCTC 204
QY 51 SerValMetGlnAspHisSerHisSerPheAspIleLysProPheThrValAspPhe 70
Db 205 TCCTGAATGAGAGACACACATCCCATCTTGATATCAAGCCCTTGACATCTGATCTTC 264
QY 71 SerSerIleSerThrProHisTyrcysAspIleProPheThrArgThrAspProVal 90
Db 265 TCCACGATTTCTACCTCCACATTCAGAGACATTCATTCACAAGACATCCAGTGGTT 324
QY 91 AlaAspTyrcysTyrcysAspLeuLysLeuGlnGlnTyrcysValAlaIleLysValGluPro 110
Db 325 GCAGATTACAGATATGACCTGAACTTCACAGAGTACCAAGTGCATATCAAGTGGAGCTT 384
QY 111 AlaSerProProTyrcysSerGlnLysThrGlnLeuTyrcys----- 124
Db 385 GCATCTCCACCTTATTTATTTGAGAGACTCAGCTCTCAATTAACCTCATGAGAGGCT 444
QY 124 ----- 124
Db 445 TCCAACTCCCTCATGATGCAATGGAATGTGCTGTGTGAGATMAAGCTTGTGATTTCAC 504
QY 124 ----- 124
Db 505 TATGAGTTCATGCTTGTGAGAGATGCAAGGGTTCTTCGAGACAAATCAGATTGAAG 564
QY 125 -----ArgAsnlyScys 128
Db 565 CTTATCTATGACAGATGATCTTAACCTGCGATCCACAAAAAGTGAATTAATGT 624
QY 129 GlnTyrcysArgphcGlnlyscysLeuAlaValGlyMetSerHisAsnAlaIleArgphe 148
Db 625 CAGTACTGTGGTTTCAGAAATGCCCTTGCAAGTGGGATGTCATATATCCATCAGGTTT 684
QY 149 GlyArgMetProGlnAlaGlnlysglyLeuLeuAlaGlnIleSerSeraspIleasp 168
Db 685 GGGCGGATGCCACAGGCCAGAGAGAGAGCTGTGGCGGAGATCTCCAGTATATCGAC 744
QY 169 GlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaValHisIleuTyrcysPhe 188
Db 745 CAGCTGATATCCAGAGTCCGCTTACCTCCGGGCCCTGGCAAAACATTTGTATGACTCATAC 804
QY 189 IleTyrcysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGlyLysThr 208
Db 805 ATAAAGTCTCTCCGCTGACCAAGCAAGGCGGAGATCTTGACAGNAAGACACA 864
```



```

Qy 209 AsplySerPropheta11etYrAspMetAsnSerLeuMetGlyGluAspLysIle 228
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 865 GACAAATC-CCATTCGGTATCTATGACATGAAATTCCTTAGATGGGAGAGATAAATC 923
Qy 229 LysPheLysHisIleThrPro 235
||||| :||| ||||| |||||
Db 924 AAGTCCACACATCACCCC 944

RESULT 12
AL543579 903 bp mRNA linear EST 16-FEB-2001
LOCUS AL543579 LTI_NFL006.PL2 Homo sapiens cDNA clone CSOD1006Y110 5
DEFINITION prime, mRNA sequence.
ACCESSION AL543579
VERSION AL543579.1 GI:12876058
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 903)
AUTHORS Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
FEATURES
source 1..903
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CSOD1006Y110"
/clone_1lb="LTI_NFL006.PL2"
/ribose_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact: Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 249 a 228 c 220 g 205 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7..71e-115 Length: 903
Score: 986.00 Matches: 200
Percent Similarity: 75.66% Conservative: 2
Best Local Similarity: 74.91% Mismatches: 3
Query Match: 28.11% Indels: 63
DB: 9 Gaps: 1
US-09-931-007a-1 (1-688) x AL543579 (1-903)
Qy 23 AlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPhetRProThraSn 42
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 97 GCCGCCGTGGCCGAGAAATGACCATGTTGACACAGAGATGCCATTCGCCACCAAC 156
Qy 43 PheGlyIleSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 157 TTTCGATATAGCTCCGTCGATCTCTCCGTAATGGAAGACCACTCCACCTCTTGTATATC 216
Qy 63 LysProPhetThrThrValAspPheSerSerIleSerThrProHisThrGluAspIlePro 82
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 217 AAGCCCTTACACTACTTCTCCACATTTCTACTCCACATTACCAAGACATTTCCA 276
Qy 83 PheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyr 102
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 277 TTCACAAGAACAGATCCAGTGGTTCAGATTAACAATATAGCTGAAGACTTCAGAGTAC 336
Qy 103 GlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGluThrGlnLeu 122
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 337 CAAGTGCAATCAAAAGTGAGCCTGCATCTCCACTTATTTATTTGAGAAGCTCAGCTC 396
Qy 123 TyrAsn----- 124
|||||
Db 397 TACATTAAGCCTCATGAAGAGCCTTCACATCCCTCATGGCAATTGATGTGTCTGT 456
Qy 124 ----- 124
Db 457 GGAGATAAGCTTCTGGATTTCATATGAGATTCATGCTTGTGAAGAGTCAAGGGTTTC 516
Qy 124 ----- 124
Db 517 TTCCGGAGAACATCATGATTAAGCTTATCTATGACAGATGTGATCTTACTGTGGATC 576
Qy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 577 CACAAAAAAGTAGAATATATGTCAGTACTGTGCGTTTCAGAAATGCCCTTCAGTGGGG 636
Qy 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLysLeuLeu 160
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 637 AAGTCTCATTAATGCCATTCAGGTTGGGCGGATGCA-CAGGCCAGAGAAGAGAGCTGTG 695
Qy 161 AlaGluIleSerSerSplLeuAspGlnLeuAsnProGlnLysAlaSplLeuArgAlaLeu 180
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 696 GCGGACATCTCCAGTATATTCAGCACGAGTAAATCCAGATCCGCTGACCTCCGGCCCTG 755
Qy 181 AlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAlaArg 200
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 756 GCAAAACATTGTATGACATCATATCAATGAATGCTCCCTACCAACAAAGCGAGAG 815
Qy 201 AlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsnSer 220
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 816 GCGATCTTGCAGGAAGACACAGCAATCATCATTCGTTATCTATGATGATTC 875
Qy 221 LeuMetGlyGluAspLys 227
||||| :||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 876 TTATATGATGGAGAGATATAA 896

RESULT 13
BO722874 872 bp mRNA linear EST 16-JUL-2002
LOCUS BO722874
DEFINITION AGENCOURT_8099880 lupsk1_sympathetic_trunk Homo sapiens cDNA clone
IMAGE:619020 5', mRNA sequence.
ACCESSION BO722874
VERSION BO722874.1 GI:21861771
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 872)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-ri@mail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM13588, row: c column: 13
High quality sequence stop: 601.
Location/Qualifiers
FEATURES
source 1..872
/organism="Homo sapiens"
/db_xref="taxon:9606"

```





OY	81	IleProPheThrArgTyrAspProValAlaIaaSPtyrLysTyrAspLeuLysLeuGln	100
Db	367	AFTCCATTCACAGAAGAACAGATCCAGTGGTGCGAGATTACAAGTAAGACTGAACTTCAA	426
OY	101	GluTyrgInSerAlaIleLysValGIuProAlaSerProProTyrrTysSerGIuLystrH	120
Db	427	GAGTAGCAAAAGTGCATCAATAAGGTGGAGCCTGCATGCCACCHTATATTTCTGAGAAAGACT	486
OY	121	GlnLeuTyrrAsn-----	124
Db	487	CAGCTCTACAAATAAGCCTCATATGAAGAGCCTTCCAACCTCCCTCATGAGCAATGAATGTGCT	546
OY	124	-----	124
Db	547	GTCCTGCGAAGATAAAGCTTCTGGAATTTCACTATGAGATTCAATCTTGTGAAGANGCAAG	606
OY	124	-----	124
Db	607	GGTTCTTCCGGAGAGAACAATCAGATTGAAGCTTATCTATGACAGATGTATCTTAATCTGT	666
OY	125	-----ArgAsnLysCysGlnTyrrCysArgPheGlnLysCysteuaLa	138
Db	667	CGGATCCACAAAAAGTAGAATAATTAAGTCACTGCTCGGTTTCAGAAAAGCCCTTGCA	726
OY	139	ValGIyMetSerHisAsnAlaIleArpPheGIyArpMetProGlnAlaGIuLysGIuLys	158
Db	727	GTTGGGAGATGTCATCAATAAGTCCATCAGGTTTTGGGCGGATGCCACAGGCCAGAGAAGGAAG	786
OY	159	LeuLeuAlaGIuLisSerSeraspIleaSpGlnLeu---AsnProGIuSerAlaaspLeu	177
Db	787	CTGTGGCGGAGATCTCCAGATGATTCGACCAGGCTGGAATCCCCAGAGTCCGCTGACCTC	846
OY	178	ArgAlaLeuAlaLysHisLeuTyrrAspSerrTyrrIleLysSerPheProLeuThr	195
Db	847	CGGGCCCTGCGCAAAACATTTGTATGACTCATCAATAAAGTCTCCGCCGCTGACA	900
<b>RESULT 15</b>			
LOCUS	AL549037	898 bp	mRNA linear EST 16-FEB-2001
DEFINITION	AL549037 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0D1042YB22 5 prime, mRNA sequence.		
ACCESSION	AL549037		
VERSION	AL549037.1 GI:12884632		
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
TITLE	Li,W.B., Gruber,C., Jesse,J. and Polayes,D.		
JOURNAL	Full-length cDNA libraries and normalization		
COMMENT	Unpublished (2001)		
	Contact: Genoscope		
	Genoscope - Centre National de Sequencage		
	BP 191 91006 Evry cedex - France		
	Email: seqref@genoscope.cns.fr, web : www.genoscope.cns.fr.		
FEATURES	Location/Qualifiers		
source	1..898		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="CS0D1042YB22"		
	/clone_lib="LTI_NFL006_PL2"		
	/tissue_type="Placenta"		
	/note="Vector: pCMVSPORT 6; Site:1: NotI; 1st strand cDNA was primed with a NotI-Oligo(gt) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifestech.com URL : http://fulllength.invitrogen.com"		

BASE COUNT	269 a	211 c	177 g	228 t	13 others
ORIGIN					
Alignment Scores:					
Pred. No.:	5.56e-107				898
Score:	924.50				190
Percent Similarity:	68.95%				1
Best Local Similarity:	68.59%				13
Query Match:	26.35%				73
DB:	9				2
				Gaps:	

US-09-931-007A-1 (1-688) x AL549037 (1-898)

QY	22	SerAlaAsnIleSerGln	-----GlnMetThr	30
		:::		
Db	58	ACCACAAATATACAAACAAGGCCATTTTCTCAACGAGAGTCAGCCTTAAACGAAATACCC		117
QY	31	MetValaSpIrrhGlnMetProPheTrpProThraSpnPhgIyIleSerSerValaSpIleu		50
Db	118	ATGGTTGACACAGAGATGCCATTCCTGGGCCACCACAACTTTGGATGCACCTCCGTGATCTC		177
QY	51	SerValMetGluAspHisSerHisSerPheAspIleIysProPheThrIrrhValaAspPhe		70
Db	178	TCCGGAATGAAGAACCACTCCACCTCTTGATATACAAAGCCCTTCACTACTGTTGACCTTC		237
QY	71	SerSerIleSerThrProHisTyrGlnAspIleProPheThrArgThrAspProValaI		90
Db	238	TCCACACATTTCTACTCCACATTTAGGAAGACATCTTCATTCACAAAGACAGATCCAGTGTT		297
QY	91	AlaAspIrrhIysTyrAspIleuIysLeuGlnGlnIrrhGlnSerAlaIleIysValaGlnPro		110
Db	298	GCAGATTCACAAGTATGACCTTAACCTTCACAGATCACCAAGATGCACATCAAGATGGACCT		357
QY	111	AlaSerProProTyrTyrSerGlnIubSrhIrnIrn	-----	124
Db	358	GCATCTCCACCTTATTATTCTGAGAAGACCTAGCTCTACATTAAGCCTCATGAAGACCT		417
QY	124	-----	-----	124
Db	418	TCCAACTCCCTCATGGCAATTGAATGTCTGTCTGTGGAGATAAAGCCTTGATTTTCAAC		477
QY	124	-----	-----	124
Db	478	TATGAGATTCATGCTTGTGAAGAGATGCAAGGGTTTCTTCGCGAGACAATCAGATTGAAG		537
QY	125	-----ArgAsnIysCys	128	
Db	538	CTTATCTATGACAGATGTGATCTTAACCTGTGCGATCCACAAAATAAGATAGAAATAATST		597
QY	129	GlnIrrhCysAlaGpHegInIubSrhCysLeuAlaIalIubSrhHisAsnAlaIleArgPhe		148
Db	598	CAGTACTCTCGTGTTCASAAATSCCTTGCACTGGGAGATGCTCTCAATTAATGCCATCAGGTTT		657
QY	149	GlyArgMetProGlnAlaGlnIubSrhIubSrhLeuAlaGlnIubSrhSerAspIleAsp		168
Db	658	GGCGGAGATGCCACAGCCCGCAAAAGCAGACAGCTTTSSCGAGATCTCAGTSATATCSAC		717
QY	169	GlnIubAsnProGlnSerAlaAspLeuArgAlaLeuAlaIubSrhIleuTyrAspSerTyr		188
Db	718	CAGCTGATCCAGAGTCCGCTGACCTCCGGCCCTCGCAAAACATTTGTATGATCTCMTAC		777
QY	189	IleIysSerPheProIubThrIlyAlaIlyAlaIrrhAlaIleIubProGlnIlySThr		208
Db	778	ATAAAGTCTTCCCGCTGACCAAAAGCAGGCGGCGATCTTGACASCAAGAGAMAMAMA		837
QY	209	AspIysSerProPheValaIleTyrAspMetAsnSerLeuMetGlnIlyu		225
Db	838	GACAATTCACCATTCGTATTCTATGACATGAATTCCTTAATGATGGAGAG		888

Search completed: February 25, 2003, 06:13:34  
Job time : 2484 secs

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 25, 2003, 04:02:41 ; Search time 273 Seconds  
(without alignments)  
1415.414 Million cell updates/sec

Title: US-09-931-007a-1  
Perfect score: 3508  
Sequence: 1 MGETLSDSPIDPSDSFTDT.....KTFEDMSLHPLQLGYKDLX 688

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 442118 seqs, 280819700 residues  
Total number of hits satisfying chosen parameters: 884236

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=framed\_p2n\_model -DRV=x1h  
-Q=/cgn2\_1/USPFO.spool/US09931007/runat\_14022003\_102554\_4718/app\_query.fasta.1.839  
-DB=Published.Applications.NA -OFMT=fastap -SUFFIX=p2n.inpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blsum62  
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR\_SCORE=pct -THR\_MAX=100  
-THR\_MIN=0 -ALIGN=15 -MODE=US09931007 -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09931007 -ECGN\_1\_1\_41 @runat\_14022003\_102554\_4718  
-NCPU=6 -ICPU=3 -NO\_XLIPY -NO\_MMAL -LARGOQUERY -NEG\_SCORES=0 -WAIT -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAEXT=7 -YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications.NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	2231	63.6	1608	US-09-765-111A-15
2	2196.5	62.6	1679	US-10-109-886-5
3	2106.5	60.0	2596	US-09-765-111A-22
4	2106.5	60.0	2625	US-09-765-111A-5

5	2099	59.8	2334	US-09-765-111A-1	Sequence 1, Appl1
6	2094	59.7	2523	US-09-765-111A-3	Sequence 3, Appl1
7	2065.5	58.9	1811	US-09-765-111A-26	Sequence 26, Appl1
8	2065.5	58.9	1811	US-09-880-107-2212	Sequence 2212, Ap
9	2065.5	58.9	2260	US-09-816-828-8	Sequence 8, Appl1
c 10	2065.5	58.9	2307	US-10-125-237-5	Sequence 5, Appl1
c 11	2065.5	58.9	2329	US-09-816-828-9	Sequence 9, Appl1
12	2065	58.9	2005	US-10-142-373-1	Sequence 1, Appl1
13	1101.5	31.4	3301	US-09-954-456-554	Sequence 554, App
14	1076.5	30.7	1323	US-10-013-807-1	Sequence 1, Appl1
15	702	20.0	428	US-10-033-528-1855	Sequence 1855, Ap
16	562	16.0	415	US-09-925-299-256	Sequence 256, App
17	512.5	14.6	611	US-09-925-299-109	Sequence 109, App
18	472	13.5	451	US-09-864-761-19039	Sequence 19039, A
19	406.5	11.6	3628	US-09-917-800A-1454	Sequence 1454, Ap
20	398.5	11.4	343	US-10-046-935-990	Sequence 990, App
21	398.5	11.4	343	US-09-878-178-990	Sequence 990, App
22	392	11.2	6282	US-10-108-605-212	Sequence 212, App
23	385.5	11.0	5970	US-10-108-605-210	Sequence 210, App
24	384	10.9	2880	US-10-044-090-361	Sequence 361, App
25	345	9.8	458	US-09-864-761-2304	Sequence 2304, Ap
26	315	9.0	2907	US-09-954-456-524	Sequence 524, App
27	313	8.9	1577	US-09-797-727-1	Sequence 1, Appl1
28	307	8.8	1679	US-10-044-090-316	Sequence 316, App
29	305	8.7	1528	US-09-962-832-109	Sequence 109, App
30	305	8.7	1528	US-09-880-107-3328	Sequence 3328, Ap
31	300	8.6	2066	US-09-895-840-1	Sequence 1, Appl1
32	281.5	8.0	1431	US-10-188-721-2	Sequence 2, Appl1
33	278	7.9	2218	US-09-880-107-3411	Sequence 3411, Ap
34	276.5	7.9	1934	US-09-909-446-1	Sequence 1, Appl1
35	276.5	7.9	1934	US-09-909-325-1	Sequence 1, Appl1
36	276.5	7.9	1934	US-09-909-326-1	Sequence 1, Appl1
37	274.5	7.8	1841	US-10-013-823-1	Sequence 1, Appl1
38	271.5	7.7	1594	US-09-964-824-559	Sequence 559, App
39	266	7.6	1752	US-09-965-703-7	Sequence 7, Appl1
40	266	7.6	1878	US-09-965-697-1	Sequence 1, Appl1
41	266	7.6	1878	US-09-965-703-6	Sequence 6, Appl1
42	266	7.6	2241	US-09-042-488B-4	Sequence 4, Appl1
43	266	7.6	2241	US-09-042-488B-6	Sequence 6, Appl1
44	266	7.6	3126	US-09-042-488B-8	Sequence 8, Appl1
45	265	7.6	1404	US-09-965-703-62	Sequence 62, Appl1

#### ALIGNMENTS

RESULT 1  
US-09-765-111A-15  
; Sequence 15, Application US/09765111A  
; Patent No. US20020106796A1  
; GENERAL INFORMATION:  
; APPLICANT: Fletcher, Jonathan A.  
; TITLE OR INVENTION: PAX8-PARAGAMMA NUCLEIC ACID MOLECULES  
; FILE REFERENCE: B0801/77196/ERP/MAT  
; CURRENT APPLICATION NUMBER: US/09/765,111A  
; CURRENT FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: US 60/177,109  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: US 60/225,079  
; PRIOR FILING DATE: 2000-08-14  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 15  
; LENGTH: 1608  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (91)...(1608)  
; US-09-765-111A-15

Alignment Scores:



Percent Similarity: 87.15% Conservative: 2  
 Best Local Similarity: 86.76% Mismatches: 2  
 Query Match: 62.61% Indels: 63  
 DB: 12 Gaps: 2

US-09-931-007a-1 (1-688) x US-10-109-886-5 (1-1679)

OY 1 MetGlgIThrLeuGlyAspSerProIleAspProGluSerAspSerPheThr 20  
 |||||  
 DB 159 ATGGGGAACCTGGGAGATTCTCTATTGACCCAGAAAGCATTCCTTCAGTATACA 218  
 OY 21 LeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPhePro 40  
 |||||  
 DB 219 CTGTGCAACATATCACAGAAATGACCATGGTTGACACAGAGATTCACATTCGGCCC 278  
 OY 41 ThrAspPheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisPhe 60  
 |||||  
 DB 279 ACCAACTTGGGATCGAGCTCGGTGATCTCTCGTAATGGAAGACCACTCCACTCTT 338  
 OY 61 AspIleLysProPheThrThrValAspPheSerIleSerThrProHisThrGluAsp 80  
 |||||  
 DB 339 GATATCAAGCCCTTCCTACTGTGTGACTTCTCCAGCATTTCTACTCCACATTCAGAGAC 398  
 OY 81 IleProPheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysGln 100  
 |||||  
 DB 399 ATTCATTCACAAGACAGATCCAGGTGGTTCAGATTACAGATGACCTGAAACTTCAA 458  
 OY 101 GluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThr 120  
 |||||  
 DB 459 GAATACCAAGTGCATCAATCAAAAGTGAGCCTGCATCCACTTATTTATTTGAGAAGACT 518  
 OY 121 GlnLeuTyrAsn----- 124  
 |||||  
 DB 519 CACCTTACAAATAGCCTCATGAGAGCCTTCACACTCCCTCATGGCAATTGATATGCT 578  
 OY 124 ----- 124  
 DB 579 GTCGTGAGATPAAAGCTTGTGATTCTACATAGAGTTCATGCTTGTGAAGATGCAAG 638  
 OY 124 ----- 124  
 DB 639 GGTTCCTCCGAGAACATCAGATTGAAGCTTATCTATAGACAGATGATCTTAATGCT 698  
 OY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138  
 |||||  
 DB 699 CGGATCCACAAAAAGTGAANTAAATGTCAGTACTGTCGGTTTCAGAAATGCCCTTGCA 758  
 OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158  
 |||||  
 DB 759 GTGGGATGTCATCATATGCCATCAGGTTGGGCGGATCGACAGGCGGAGAAAGAGAG 818  
 OY 159 LeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGluAlaAspLeuArg 178  
 |||||  
 DB 819 CTGTGTGGCGAGATCTCCAGTATATCGACAGCTCAATCCAGAGCCCTCACTCCCT 878  
 OY 179 ---AlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197  
 |||||  
 DB 879 CAGGCCCTGGCAAAACATTTGATATGCTCATACATATAATCTCTCCCGCTGACCAAGCA 938  
 OY 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217  
 |||||  
 DB 939 AAGGCGAGGCGCATCTTGCAGAGAAAGACACAGACAAATCAACATTCCTTATCATATGAC 998  
 OY 218 MetAsnSerLeuMetMetLysGluAspLysIleLysPheLysHisIleThrProLeuGln 237  
 |||||  
 DB 999 ATGAATTCCTTATATGATGGAGAGATAAATCAATCAATCAACACTCAACCCCTCGCAG 1058  
 OY 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257  
 |||||  
 DB 1059 GAGCAACAGCAAGAGTGGCCATCCCATCTTTCAGGGCTGCCAGTTTGGCTCCGGGAG 1118  
 OY 258 AlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277  
 |||||

DB 1119 CCTGTGACAGAGATCACAGACTATGTCAAAAGCATTCCTGTTTGTAAATCTGACTTG 1178  
 OY 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrLeuAla 297  
 |||||  
 DB 1179 AAGGACCAAGTACTCTCTCAAAATATGGATGTCACAGATCATTTTACCAATGCTGGCC 1238  
 OY 298 SerLeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThrArgGlu 317  
 |||||  
 DB 1239 TCCTTGATGATAAATGAATGGGTTCTCATATCCGAGGGCCAAAGCTTCATGCAAGGGCAG 1298  
 OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337  
 |||||  
 DB 1299 TTCTTAAAGAGCTCGGAAGCCTTTGGTGTACTTTATGAGACCCCAAGTTTGAAGTTGCT 1358  
 OY 338 ValLysPheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIle 357  
 |||||  
 DB 1359 GTGAAGTTCAATGACCTGGAAATGATGACACAGCATGGGCAATTTATTTGCTGTCAAT 1418  
 OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377  
 |||||  
 DB 1419 ATTCACAGTGGAGACCGCCGAGGTTTGTGATGTGGAAGCCCATTTGAACACATTCAAGAC 1478  
 OY 378 AsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeu 397  
 |||||  
 DB 1479 AACCTGCTACAGCCCTGGAGCTCCAGCTGAGCTGAAGCTGAACCATCTGAGCTCCACAGCTG 1538  
 OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGln 417  
 |||||  
 DB 1539 TTTGGCAACCTCTCCAGAAATGACAGACCTCAACACAAATTTGTACGGAACACGCTGAG 1598  
 OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437  
 |||||  
 DB 1599 CTACTGCAAGTATATCAAGAGACGAGACAGACATGAGTCTTACCCGCTCTGCAAGAG 1658  
 OY 438 IleTyrLysAspLeuTyr 443  
 |||||  
 DB 1659 ATCTACAAGACTGTGAC 1676

# RESULT 3

US-09-765-111A-22  
 ; Sequence 22, Application US/09765111A  
 ; Patent No. US20020106796A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Fletcher, Jonathan A.  
 ; APPLICANT: Kroll, Todd G.  
 ; TITLE OF INVENTION: PAX8--PARAGAMMA NUCLEIC ACID MOLECULES  
 ; FILE REFERENCE: B0801/7196/ERP/MAT  
 ; CURRENT APPLICATION NUMBER: US/09/765,111A  
 ; CURRENT FILING DATE: 2001-01-18  
 ; PRIOR APPLICATION NUMBER: US 60/177,109  
 ; PRIOR FILING DATE: 2000-01-20  
 ; PRIOR APPLICATION NUMBER: US 60/225,079  
 ; PRIOR FILING DATE: 2000-08-14  
 ; NUMBER OF SEQ ID NOS: 47  
 ; SOFTWARE: FastSeq for Windows Version 3.0  
 ; SEQ ID NO 22  
 ; LENGTH: 2596  
 ; TYPE: DNA  
 ; ORGANISM: Homo Sapiens  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (161)...(2596)  
 ; US-09-765-111A-22

## Alignment Scores:

Pred. No.: 2,96e-237 Length: 2596  
 Score: 2106.50 Matches: 425  
 Percent Similarity: 84.45% Conservative: 4  
 Best Local Similarity: 83.66% Mismatches: 14  
 Query Match: 60.05% Indels: 65  
 DB: 10 Gaps: 3



```

QY 18 ThrAspThrLeuSerAlaAsnIleSerGlnGluMetThrMetValAspThrGluMetPro 37
   ::: ::::
Db 1159 GCCTCTCTGCCATCGCAGCGCATGTGGCAGAAATGACCATGTGGACACAGATGCCA 1218
QY 38 PheTTPProThrAsnPhcGlyIleSerSerValAspLeuSerValMetGluAspHisSer 57
   |||||
Db 1219 TTCTGGCCCAACACTTGGATCGAGTCGCCGATCTCTCCGTATGTGAACACCATCC 1278
QY 58 HisSerPheAspIleLeuSerProPheThrValAspPheSerSerIleSerThrProHis 77
   |||||
Db 1279 CACTCTTTGATATCAAGCCCTCACTACTGTGACTCTCCAGCATTTCTACTCCACAT 1338
QY 78 TyrGluAspIleProPheThrArgThrAspProValAlaAspTyrIleTyrAspLeu 97
   |||||
Db 1339 TACGAAGACATTCATTCACAGAAGATCCAGTGTTGCCAGATTACAGATATGACCTG 1398
QY 98 LysLeuGlnGluTyrGlnSerAlaIleLysValGluProAlaSerProProTyrTyrSer 117
   |||||
Db 1399 AACCTTCAAGAGTACCAAAAGTGCATCAAAAGGAGCCGTCATCCACCTATTATCT 1458
QY 118 GluLysThrGlnLeuTyrAsn----- 124
   |||||
Db 1459 GAGAAGACTCACCTCTACAAATAAAGCCTCATGAAGACCTTCCACTCCCTCATGCAAT 1518
QY 124 ----- 124
Db 1519 GAATGTCGTGTGTGGAGATAAAGCTTCTGATTCACATGAGATTGATCTTGTGAA 1578
QY 124 ----- 124
Db 1579 GGATGCAAGGTTTCTTCCGGAGAACATCAGATTGAAGCTTATCATGACAGATGTGAT 1638
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLys 135
   |||||
Db 1639 CTTAACTGTGGATCCACAAAAAAGTAAATAATGATGATGCTGCGGTTTACAGAAA 1698
QY 136 CysLeuAlaValAlaGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnLaglu 155
   |||||
Db 1699 TGCCTTGCAGTGGGGATGCTCATATGATGCATCAGTTGGGCGGATGCCACAGGCCGAG 1758
QY 156 LysGluLysLeuLeuAlaGluIleSerSerAspIleAspGlnLeuAsnProGlnSerAla 175
   |||||
Db 1759 AAGGAGAAGCTGTGGCGAGATCTCCAGTGAATGACACCACTGAATCCAGATCCGCT 1818
QY 176 AspLeuArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThr 195
   |||||
Db 1819 GACCTCCGGGCCCTGGCAAAACATTTGTATGACTCATACATAAAGTCTCCCGCTGAC 1878
QY 196 LysAlaLysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIle 215
   |||||
Db 1879 AAAGCAAAAGCGAGGCGCATCTTGACAGAAAGACACAGCAAAATCCACATTCGTTATC 1938
QY 216 TyrAspMetAsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrPro 235
   |||||
Db 1939 TATGACATGAATTCCTTATGATGAGAGATTAATAATCAATGTCAAACACATCACCCCC 1998
QY 236 LeuGlnGlnGlnSerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSer 255
   |||||
Db 1999 CTGCGAGGACACAGCAAAAGAGTGGCCATCCGATCTTTCAGAGGCTGCCAGTTGCTCC 2058
QY 256 ValGluAlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeu 275
   |||||
Db 2059 GTGGAGGCGTGCAGAGATCACAGATATGCCAAAGACATTCCTGGTTTATAATCTT 2118
QY 276 AspLeuAsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMet 295
   |||||
Db 2119 GACTTGAAGACAGCAAGTAACTCTCTCAAAATATGAGATCCAGATCATTTACACAAT 2178
QY 296 LeuAlaSerLeuMetAsnLysAspGlyValLeuIleSerGlnGlyGlnGlyPheMetThr 315
   |||||
Db 2179 CTGGCTCTCTTATGAATAAAGATGGGTTCTCATATCCGAGGCGCAAGGCTTCATGACA 2238

```

```

QY 316 ArgGluPheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGlu 335
   |||||
Db 2239 AGGAGTTTCTTAAGACGCTCGGAAGCCTTTGGTGAATTTATGAGCCCAAGTTTGAC 2298
QY 336 PheAlaValLysPheAsnAlaLeuGlnLeuAsnAspSerAspLeuAlaIlePheIleAla 355
   |||||
Db 2299 TTTCGTGTGAAGTCTCAATGCACCTGGAATATGATGACAGGACTTGCCATATTTATCT 2358
QY 356 ValIleIleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIle 375
   |||||
Db 2359 GTCATTTATTTCTAGTGGAGACCGCCAGGTTGCTGAAATGTGAAGCCCATTTGAAGACAT 2418
QY 376 GlnAspAsnLeuLeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGlnSerSer 395
   |||||
Db 2419 CAGACAACTCTCTACAGCCCTGGAGCTCCAGCTGGAAGCTAACACCCCTGAGTCTCA 2478
QY 396 GlnLeuPheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHis 415
   |||||
Db 2479 CAGCTGTTTGGCCAACTGCTCCGAAATGACAGACCTCAGACAGATTGTCACGGAGAC 2538
QY 416 ValGlnLeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeu 435
   |||||
Db 2539 GTGCAGCTACTGCAAGTATCAAGAAAGGAGACAGACATGATGTTTACCCGCTCTG 2598
QY 436 GlnGluIleTyrLysAspLeuTyr 443
   |||||
Db 2599 CAGGAGATCTACAGAGACTTTGAC 2622

RESULT 5
US-09-765-111A-1
: Sequence 1, Application US/09765111A
: Patent No. US20020106796A1
: GENERAL INFORMATION:
: APPLICANT: Fletcher, Jonathan A.
: TITLE OF INVENTION: PAX8-PPARGAMMA NUCLEIC ACID MOLECULES
: FILE REFERENCE: BOB01/77196/ERP/MAT
: CURRENT APPLICATION NUMBER: US/09/765.111A
: CURRENT FILING DATE: 2001-01-18
: PRIOR APPLICATION NUMBER: US 60/177,109
: PRIOR FILING DATE: 2000-01-20
: PRIOR APPLICATION NUMBER: US 60/225,079
: PRIOR FILING DATE: 2000-08-14
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 1
: LENGTH: 2334
: TYPE: DNA
: ORGANISM: Homo Sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(2334)
: US-09-765-111A-1

Alignment Scores:
Pred. No.: 1.91e-236 Length: 2334
Score: 2099.00 Matches: 420
Percent Similarity: 84.33% Conservative: 5
Best local Similarity: 83.33% Mismatches: 13
Query Match: 59.83% Indels: 66
DB: 10 Gaps: 2

US-09-931-007A-1 (1-688) x US-09-765-111A-1 (1-2334)
QY 6 GlyAspSerProIleAspProGlnSerAspSerPheThrAspThrLeuSerAlaAsn--- 24
   ||| :::: |||||
Db 820 GGGAGGCGACCCCTGACCCCTTCCAAACGCCACATGGGGCGCAACCTCTCAGCTACACG 879
QY 25 -----IleSerGlnGluMetThrMetValAspThrGluMetProPheThrProThr 41
   |||||
Db 880 ACCTACCCCGTGTGGCAGAAATGACATGTGTGACACAGAGATGCCATTCCTGGCCGACC 939

```



QY	42	AsnPhgEgLIIEserSerValAspLeuSerValMetGlnAspHisSerHisSerPhasp	61
Db	940	AACTTGGGATACGTCGGTGAGATCTCCGTATGGAAGACACCCACCTCCTTGTAT	999
QY	62	IlElvsProPheThrThrValAspPheSerSerLIEserThrProHisTyrGlnAspIle	81
Db	1000	ATCAAGCCCTTCACTACGCTTGACTTCTCCAGCATTTTCACTCCACATTAGCAACACTT	105
QY	82	ProPheThrArgThrAspProValValAlaAspTyrLysTyrAspLeuLysLeuGlnGlu	101
Db	1060	CCATTCAACAAGAACAGATCCAGTGGTTCAGATTACAAGATGATGACTGCAACTTCAGAG	1119
QY	102	TyrGlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerLulysThrGln	121
Db	1120	TACCAAAATGCAATCAAAAGTGAGCCGTGCATCTCCACCTTATTTATCTGAGAAGACTGAG	1179
QY	122	LeuTyrAsn	124
Db	1180	CTCTACATTAAGCCTCATGAAGACCTTCCAACTCCCTCATGGCAATTGAATGCTGTGTC	1238
QY	124		124
Db	1240	TGTGAGATAAAGCTTCGTGATTTCACTATGAGATTTCATGCTTGAAGATGACAAGGT	1299
QY	124		124
Db	1300	TTCTCCGGAGACAATCAGATTGAACCTTATCTATGACAGATGATGATCTTAACGTCCG	1359
QY	125	-----ArgasnlvsCysGlnTyrCysArgPheGlnLysCysLeuAlaVal	139
Db	1360	ATCCACAAAAAAGTACGAATTAATATGTCAGTACTGTCGGTTTCAGAAATGCTTCCAGTG	1419
QY	140	GlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeu	159
Db	1420	GGGATGCTCATTAATGCCATCAGGTTTGGCGCGATGCCACAGGCGAGAGAAGAACTCG	1479
QY	160	LeuAlaGlnLIEserSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArgAla	179
Db	1480	TTGGCGGAGATCTCCAGTGAATTCGACACACTAAATCCAGAGTCCGTACCTCCGGGCC	1539
QY	180	LeuAlaIysHisLIEuTyrAspSerTyrLIElysSerPheProLeuThrLysAlaLysAla	199
Db	1540	CTGGCAAAAACATTGTGTATGACTCTATCATATAAATCTCTCCCGCTGACCAAGCAAGGCG	1599
QY	200	ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIIEtyrAspMetAsn	219
Db	1600	AGGCGATCTTGACAGGAAGACAAACAGACAATCAACATCGTTATCTATCATCATGAT	1659
QY	220	SerLeuMetMetGlyGlnAspLysLIElysPheLysHisLIEthrProLeuGlnGlnGln	239
Db	1660	TCTCTTAATGATGGAGAAAGATAAATCAAGTTCAAAACATCAACCCCTCGCAGAGACAG	1719
QY	240	SerLysGlnValAlaIleArgGlyIlePheGlnGlyCysGlnPheArgSerValGlnAlaVal	259
Db	1720	AGCAAAAGAGTGGCCATCCGCACTTTTCAGGGCTGCAGATTGGCTCCGTGGAGGCTGTG	1779
QY	260	GlnGlnLIEthrGlnTyrAlaLysSerLIEProGlnLysPheValAsnLeuAspLeuAsnAsp	279
Db	1780	CAGGAGATCACAGATGATGCCAAAGACATTCCTGGTTTGTAAATCTTGACTTGAGACAC	1839
QY	280	GlnValThrLeuLeuLysTyrGlyValHisGlnIleIIEtyrThrMetLeuAlaSerLeu	299
Db	1840	CAAGTAATCTCTCCCAATATATGAGATCCAGAGATCATTTACACATGCTGGCTCTCTTG	1899
QY	300	MetAsnLysAspGlyValLIEuLIEserGlnGlyGlnGlyPheMetThrArgGlnPheLeu	319
Db	1900	ATGATATAAAGTGGGTTCATATCCGAGGCGCAAGGCTTATATACAAGGAGATTTCTTA	1959
QY	320	LysSerLeuArgLysProPheGlyAspPheMetGlnProLysPheGlnPheAlaValLys	339
Db	1960	AAGAGCCTCGCAAAAGCTTTTGGTGACTTTATGAGACCCAAATTTGACTTTGCTGTGAG	2019
QY	340	PheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIIEpheIIEAlaValIIEIleLeu	359

Db	2020	TTCAATGCACCTGGAAATTAGATGACAGCGACTTGGCAATATTATTGCTCATTTTCTC	20799
Qy	360	SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu	379
Db	2080	AGTGAGAGACCCGCCAGGTTTGCTGAATGTGAACCCCATTTGAAGACATTCACAGCAACTG	21399
Qy	380	LeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla	399
Db	2140	CTACAGCGCCCTGGAGGCTCCAGCTGAACCTGAACCCCTGAGTCTCCACAGCTGTGTTGCC	21999
Qy	400	LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeu	419
Db	2200	AAGCTGCTCCCGAAATGACAGACGCTTCAGACAGATTGTACGGACACCTGAGGCTACTG	22599
Qy	420	GlnValIleLysLysThrGlnThrAspMetSerLeuHisProLeuLeuGlnGluIleLeuYr	439
Db	2260	CAGGTGATCAAGAAAGACGAGGAGACAGACATGAGTCTTACCCGCTCCTGAGAGATGATTAC	23199
Qy	440	LysAspLeuYr 443	
Db	2320	AAGGACTTGTAC 2331	

```

RESULT 6
US-09-765-111A-3
; Sequence 3, Application US/09765111A
; Patent No. US20020106796A1
; GENERAL INFORMATION:
; APPLICANT: Fletcher, Jonathan A.
; APPLICANT: Kroll, Todd G.
; TITLE OF INVENTION: PAX8-PPARGamma NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: AND POLYPEPTIDES AND USES THEREOF
; FILE REFERENCE: B0801/7196/ERP/MAT
; CURRENT APPLICATION NUMBER: US/09/765,111A
; PRIOR FILING DATE: 2001-01-18
; PRIOR APPLICATION NUMBER: US 60/177,109
; PRIOR FILING DATE: 2000-01-20
; PRIOR APPLICATION NUMBER: US 60/225,079
; PRIOR FILING DATE: 2000-08-14
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 2523
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1) ... (2523)
US-09-765-111A-3

Alignment Scores:
Pred. No.:      8.31e-236      Length:      2523
Score:          2094.00      Matches:      416
Percent Similarity: 87.03%      Conservative: 0
Best Local Similarity: 87.03%      Mismatches: 0
Query Match:    59.69%      Indels:      62
DB:             10          Gaps:         1

US-09-931-007A-1 (1-688) x US-09-765-111A-3 (1-2523)
QY      28  GUMETThrmetValAspThrGlumetProPheThrProThraSnphagilyIleSerSer  47
      |||||||
Db      1087  GAAATGACCAATGGTTGACACACAGAGATGGCATTCGGCCACCACTTGGGATCAGCTCC  11
QY      48  ValAspIleSerValMetGluAspHisSerHisSerPheAspIleIysProPheThrThr  67
      |||||||
Db      1147  GTGGATCTCTCCGTAATGGAGACACACCCACCTCTTGTATACAGACCCCTTCACACT  120
QY      68  ValAspPheSerSerIleSerThrProHisIyrcIuAspIlePheProPheThrArgThrAsp  87
      |||||||
Db      1207  GTTGACCTTCTCCAGACTTTTACTCTCCACTTACGATTCACATTCACATTCACAAAGACAGAT  120
QY      88  ProValValAlaAspTyrlYsTyrAspIleuIysLeuGlnGluIyrcIuSerAlaIleIys  107

```



|||||  
Db 1267 CCAGGTGTTGCAGATTACAGATGACCTGAACTTCAAGATACCAAGATGCAATCAAA 1326  
OY 108 ValG1ProL1aserProProTyTySerGluLysThcGluLeuTyAsn ----- 124  
Db 1327 GTGAGGCTGTCATCTCCACCTTATTATTCTGAGAGAGCTCAGCTTACATTAAGCTCAT 1386  
OY 124 ----- 124  
Db 1387 GAGAGGCTTCCAACTCCCTCATGSCAATTGATGTGCTGTCTGTGAGATAAAGCTTCT 1446  
OY 124 ----- 124  
Db 1447 GGATTTCACATGAGATGATGCTGTGAGAGATGACAGGCTTCTTCGGAGAACATC 1506  
OY 125 ----- Arg 125  
Db 1507 AGATTGAGCTTCTATGACAGATGATGATCTTAAGTCCGATCCCAAAAAAGTAGA 1566  
OY 126 AsnLysCysGlnTyrcysArgPheGlnLysCysLeuAlaValG1LysMetSerHisAsnAla 145  
Db 1567 AATAAATGTCAGTACTGTGGCTTTCAGAAATGCTTGACAGTGGGAGATGCTCATTAATGCC 1626  
OY 146 L1earGPhcG1yArgMetProGlnAlaGluLysGluLysLeuAlaGluL1LysSerSer 165  
Db 1627 ATCAGGTTTGGGGGATGCGACAGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1686  
OY 166 AspL1eAspGlnLeuAsnProGlnSerAlaAspLeuArgAlaLeuAlaLysHisL1eUtyr 185  
Db 1687 GATATGACACAGTGAATCCAGAGTCCGCTGACCTCCGGGCTTGGCAAAACATTTGTAT 1746  
OY 186 AspSerTyrl1eLysSerPheProLeuThrLysAlaLysAlaArgAlaL1eLeuThrGly 205  
Db 1747 GACTCATACATAAAGTCCTCCCGTGACCAAGCAAGGCGGAGATCTTGACAGGA 1806  
OY 206 LysThrThrAspLysSerProPheVal1LeyTyAspMetCAsnSerLeuMetGlyGlu 225  
Db 1807 AAGACAAACAGCAAAATCACCATTCTGATCATGATGATGATGATGATGATGATGATG 1866  
OY 226 AspLysL1eLysPheLysHisL1eThrProLeuGlnGlnL1eSerLysGluValAlaL1e 245  
Db 1867 GATTAATACAGTTCAAACATCACATCCCCCTCAGAGAGAGAGCAAGAGGGGCGCATC 1926  
OY 246 ArgL1ePheGlnGlnCysGlnPheArgSerValGluAlaValGlnGlnL1eThrGluTy 265  
Db 1927 CGCATCTTTCAGGGGCGCAGCTTCCCTCGTGAGGCTGTGAGGAGATCACAAGATTA 1986  
OY 266 AlaLysSerL1eProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285  
Db 1987 GCCAAAAGCATTCCTGTTTGTAAATCTGACTGAACAGCAACCACTTAACCTCCCTCAA 2046  
OY 286 TyrgL1ValHisGlnL1eLeyTyThrMetLeuAlaSerLeuMetLysAspGlyVal 305  
Db 2047 TATGAGTCCACAGATCATTTTACACAATGCTGGCTCTTGATGATTAAGATGGGGTT 2106  
OY 306 LeuL1eSerGlnGlnGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325  
Db 2107 CTCATATCCGAGGGCCAAAGGCTTCATGACAAAGGAGATTCTTAAGAAGCCCTGCAAG 2166  
OY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345  
Db 2167 TTTGGTACTTATGAGGCCCAAGTTGAGTTGCTGTGAAGTTCATACACATGCAATTA 2226  
OY 346 AspAspSerAspLeuAlaL1ePheL1eAlaValL1eL1eLeuSerGlyAspArgProGly 365  
Db 2227 GATGACAGCGACCTTGCAATATTATTTGCTGCTCATTTATCTCAGAGACCCCAAGGT 2286  
OY 366 LeuLeuAsnValLysProL1eGlnAspL1eGlnAspAsnLeuLeuGlnAlaLeuGluLeu 385  
Db 2287 TTTCTGAATGTGAAGCCCATTTGAAGACATTCAGACCAACCTGCTACAAAGCCCTGAG 2446  
OY 386 GlnLeuLysLeuAsnHisProGlnSerGlnLeuPheAlaLysLeuLeuGlnLysMet 405  
|||||

Db 2347 CAGCTGAAGCTGAACCAACCCCTGAGTCCATCAGCTGTTTGCCAAAGCTGCTCCAGAAATG 2406  
OY 406 ThrAspLeuArgGlnL1eValThrGluHisValGlnLeuLeuGlnValL1eLysLysThr 425  
Db 2407 ACAGACCTCAGACAGATTTGTACAGGAGACACGTCTACACTACTGAGGTGATCAAGAGAGAG 2466  
OY 426 GluThrAspMetSerLeuHisProLeuLeuGlnGlnL1eTyrcysAspLeuTy 443  
Db 2467 GAGACAGACATGATGCTTCCACCCGCTCCTGACAGAGATCTACAAAGACATTTGTAC 2520  
RESULT 7  
US-09-765-111A-26  
; Sequence 26, Application US/09765111A  
; Patent No. US20020106796A1  
; GENERAL INFORMATION:  
; APPLICANT: Fletcher, Jonathan A.  
; APPLICANT: Kroll, Todd G.  
; TITLE OF INVENTION: PAX8-PERAGRAMMA NUCLEIC ACID MOLECULES  
; FILE REFERENCE: B0801/7196/ERP/MAT  
; CURRENT APPLICATION NUMBER: US/09/765,111A  
; PRIOR FILING DATE: 2000-01-18  
; PRIOR APPLICATION NUMBER: US 60/177,109  
; PRIOR FILING DATE: 2000-01-20  
; PRIOR APPLICATION NUMBER: US 60/225,079  
; PRIOR FILING DATE: 2000-08-14  
; NUMBER OF SEQ ID NOS: 47  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 26  
; LENGTH: 1811  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (173)..(1609)  
US-09-765-111A-26  
Alignment Scores:  
Pred. No.: 1,11e-232 Length: 1811  
Score: 2065.50 Matches: 414  
Percent Similarity: 85.71% Conservative: 6  
Best Local Similarity: 84.49% Mismatches: 7  
Query Match: 58.88% Indels: 63  
DB: 10 Gaps: 2  
US-09-931-007A-1 (1-688) x US-09-765-111A-26 (1-1811)  
OY 23 AlaAsnL1eSerGlnGlnL1eThrMetValAspThrGluMetProPheThrProThrAsn 42  
Db 155 GCCGCGGTGGCCGCAAGAAATGACATGATTGACACAGAGATGCGATTCTGGCCACCAAC 214  
OY 43 PheGlyL1eSerSerValAspLeuSerValMetGluAspHisSerLysPheAspL1e 62  
Db 215 TTTGGATGATGCTCCGTGATCTCTCGTAATGGAAGACACATCCCATCTTGTGATATC 274  
OY 63 LysProPheThrThrValAspPheSerL1eSerThrProHisTyrcysAspL1ePro 82  
Db 275 AAGCCCTTACATCTGTTGACTTCTCCAGCATTTCTACTCCACATTAAGACATTCGA 334  
OY 83 PheThrArgThrAspProValValAlaAspTyrcysLysTyrcysAspLeuLysLeuGlnGlu 102  
Db 335 TTCACAAAGAACAGATCCAGTGTTCAGATTCACAAATAGTACCTGAAACTTCAAGAGTAC 394  
OY 103 GlnSerAlaL1eLysValGluProAlaSerProProTyrcysSerGlnLysThrGlnLeu 122  
Db 395 CAAGTGCATCAATAAGTGGAGCTGATCTCCACCTTATTATTCTGAGAAAGACTCAGCTC 454  
OY 123 Tyrcys ----- 124  
Db 455 TACAATAACCTTCATGAAGAGCTTCCAACTCCCTCATGGAATTAATGCTGTCTCTGT 514  
OY 124 ----- 124

```
Db 515 GGAGATAAAGCTTCTGGATTTCATGATGAGTTTCATGCTTGTGAAGATGCAAGGTTTC 574
QY 124 ----- 124
Db 575 TTCCGGAGAACATCAGATTGAACCTTATCTANGACAGATGTGATCTTAAGTGTGGATC 634
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
Db 635 CACAAAAAAGTGAATAATTAATGCTACTGCTGCTTTCAGAAATGCTTGGAGTGGG 694
QY 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLysLeu 160
Db 695 ATGCTCATATGCTCATCAGATTGGCGGTGCTCCACAGCGCGGAGAGAGAACTGTTG 754
QY 161 AlAGlnLysSerSerAspIleAspGlnLeuAsnProGlnSerAlaAspLeuArg---Ala 179
Db 755 GCGGAGATCTCCAGTATATGACCACTGATCAGAGTCCAGAGTCCGCTCCGACAGCC 814
QY 180 LeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
Db 815 CTGGCAAAACATTTGTATGACTCATACATAAAGTCCCTTCCCGCTGACCAAGCAAGCGC 874
QY 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
Db 875 AGGCGCATCTTGAACGAGAAAGACAAACAGCAAAATCACCATTCCTTATCTATGACATGAAT 934
QY 220 SerLeuMetMetGlyGlnLysLysIleLysPheLysHisIleThrProLeuGlnGlnGln 239
Db 935 TCCCTTAATGATGGAGAGATTAATAATCAAGTCAAAACATCACACCCCTGACAGAGACAG 994
QY 240 SerLysGlnValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlnAlaVal 259
Db 995 AGCAAAAGGTGGCCATCCGCACTTTTCAGAGGCTGCCAGTTGCTCCGTGAGAGCTGTG 1054
QY 260 GlnGlnLysLeuThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAsp 279
Db 1055 CAGGAGATCACAGAGTATGCCAAAGCATTCCTGTTTGTAAATCTTGAAGTTCGAGACGAC 1114
QY 280 GlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeu 299
Db 1115 CAGATCACTCTCCCAATATGAGTCCAGAGATCACTTTCACAAATGCTGGCCCTCTTG 1174
QY 300 MetAsnLysAspGlyValLeuLysSerGlnGlyGlnLysPheMetThrArgGlnPheLeu 319
Db 1175 ATCAATTAAGATGGGGTTCATATCCGAGGCGCAAGGCTTCATGCAAGAGGATTTCTA 1234
QY 320 LysSerLeuArgLysProPheGlyAspPheMetGlnProLysPheGlnPheAlaValLys 339
Db 1235 AAGAGCCTGGAAAGCCTTTTGGTGAATTCATGAGGCCCAAGTTTGAGTTGCTGTGAG 1294
QY 340 PheAsnAlaLeuGlnLeuAspAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 1295 TTCAATGCTCATGAAATTAATGACAGGACTTGGCAATATTATTGCTGCTCATATTATCTC 1354
QY 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGlnAspIleGlnAspAsnLeu 379
Db 1355 AGTGAAGACCGCCAGGTTTGTGCAATGTGAACCCCAATTGAAGACATTCAGAACAACTCG 1414
QY 380 LeuGlnAlaLeuGlnLeuGlnLysLeuAsnHisProGlnSerSerGlnLeuPheAla 399
Db 1415 CTACAGAGCCCTGGAGCTCCCGCTGAAGCTGAACACCCTGAGTCCACAGCTGTTGGCC 1474
QY 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGlnHisValGlnLeuLeu 419
Db 1475 AACCTGCTCCAGAAATGACAGACTCAGACAAATTTGTCAGGAACACGTCGAGCTACTG 1534
QY 420 GlnValIleLysLysThrGlnTyrAspMetSerLeuHisProLeuGlnGlnLysIleTyr 439
Db 1535 CAGGTGATCAAGAGAGGAGACAGACATGAGTCTTACCCGCTCCGTCGAGAGATCTTAC 1594
QY 440 LysAspLeuTyrAlaThrAlaIleLeuThr 449
Db 1595 AAGGACTTGTACTAGCAGAGAGTCTGTGAGC 1624
```

```
RESULT 8
US-09-880-107-2212
; Sequence 2212, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2212
; LENGTH: 1811
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 LA0904
US-09-880-107-2212

Alignment Scores:
Pred. No.: 1,11e-232 Length: 1811
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2

US-09-931-007a-1 (1-688) x US-09-880-107-2212 (1-1811)
QY 23 AlaAsnLysSerGlnGlnMetThrMetValAspThrGlnMetProPheThrProThrAsn 42
Db 155 GCGCGCTGGCCGCAAGAAATGACATGTTGACACAGAGATCCATCTTGGCCACCAC 214
QY 43 PheGlyLysSerSerValAspLeuSerValMetGlnAspHisSerHisSerPheAspIle 62
Db 215 TTGGGATGACGTCCTGGGATCTCTCGTATATGGAAGACACATCCCATCTTGTGATATC 274
QY 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGlnAspIlePro 82
Db 275 AAGCCCTTCACTAGTGTGACTTCTCCAGCATTTCTACATCCACATTAACGAAGACATTC 334
QY 83 PheThrArgThrAspProValAlaIleAspTyrLysTyrAspLeuLysLeuGlnLysTyr 102
Db 335 TTCAACAAGACAGATCCAGTGGTTGCAAGATTACAAAGTATGACCTGGAACCTCAAGGTAC 394
QY 103 GlnSerAlaIleLysValGlnProAlaSerProProTyrTyrSerGlnLysTyrThrGlnLeu 122
Db 395 CAAGAGTCATCAATAAGTGAAGCTGCATCTCCACCTTATATTCTGAGAAAGATCAGCTC 454
QY 123 TyrAsn----- 124
Db 455 TACAATAAGCCTCATGAAGAGCCTTCCAACTCCCTCATGCAATGGAATGTGTCGTGT 514
QY 124 ----- 124
Db 515 GGAGATAAAGCTTCTGGATTTCATGATGAGTTCATGCTTGTGAAGAGTCAAGGTTTC 574
QY 124 ----- 124
Db 575 TTCCGGAGAACATCAGATTGAAGCTTATCTATGACAGATGTGATCTTAAGTGTGGATC 634
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
Db 635 CACAAAAAAGTGAATAATTAATGCTACTGCTGCTTTCAGAAATGCTTGGAGTGGG 694
```



```
Db 1177 CACAAAAAGTAGAATAAATGTCTAGTCTGCGTTTCAGAAATGCTTCGAGTGGG 1118
QY 141 MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluArgLysLeu 160
Db 1117 ATGTCTCATTAATGCCATCAGGTTTGGCGGATGCGACAGGCCAGAGAGGAGGAGCTGTG 1058
QY 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg---Ala 179
Db 1057 GCGGAGATCTCCAGTGATATCGACAGCTGAATCGAGAGTCCGCTGACCTCCGTCAGGCC 998
QY 180 LeuAlaLeuHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199
Db 997 CTGGCAAAACATTGTGTATGACTCATATCAATAAAGTCTCCGCTGACCAAGCAAGGCG 938
QY 200 ArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAspMetAsn 219
Db 937 AGGGCATCTTGACAGGAAGAAGACACAGACAAATCCATTGCTTATCTATGACATGAAT 878
QY 220 SerLeuMetMetIlgLysLysLysIleLysPheLysHisIleThrProLeuGlnGluGln 239
Db 877 TCCTTAATGATGGAGAAATAAATCAAGTTCAACACATCACCCCTCGCAGAGCAG 818
QY 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259
Db 817 ACCAAAGAGGTGGCCATCCGCACTTTTCAGGGCTGCCAGTTTCCGCTGGAGGCTGTG 758
QY 260 GlnGluIleThrGluTyrAlaLysSerIleProGluPheValAsnLeuAspLeuAsnAsp 279
Db 757 CAGGAGATCACAGAGATATCCAAANAGCATTCCTGGTTTGTAAATCTTGACTTGAAGCAG 698
QY 280 GlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299
Db 697 CAAGTAACCTCTCTCAAAATATGAGTCCACAGATCATTTACACATATGCTGGCCCTCTTG 638
QY 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgLysPheLeu 319
Db 637 ATGAATAAAGATGGGTCTCATATCCGAGGGCCAGGCTTCATGTCAAGGAGTTTCTA 578
QY 320 LysSerLeuArgLysProPheGlnAspPheMetGluProLysPheGluPheAlaValLys 339
Db 577 AAGAGCTTCGGAAGCCCTTTGGTGTGACTTATGAGAGCCCAAGTTTGAAGTTGCTGTAAG 518
QY 340 PheAsnAlaLeuGlnLeuAspSerSerAspLeuAlaIlePheIleAlaValIleIleLeu 359
Db 517 TTCAAAGCACTGGAATTAATGATGACACGCACTGGCAATATTATATCTGCTCATATTATTC 458
QY 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379
Db 457 AGTGGAGACCGCCAGGTTTGTGATGTGAAGCCCATTTGAAGACATTCAGAACCAACCTG 398
QY 380 LeuGlnAlaLeuGlnLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399
Db 397 CTACCAAGCCCTGGAGCTTCAGCTGAAGCTGAACCAACCCCTGAGTCCCTACAGCTGTGGCC 338
QY 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGlnLeuLeu 419
Db 337 AAGCTCTCTCCAAATAATGACAGACCTCAGACAGATTTGTACAGGAACACCTGCAGCTACTG 278
QY 420 GlnValIleLysLysThrGlnThrLysPheMetSerLeuHisProLeuLeuGlnGluIleTyr 439
Db 277 CAGGTGATCAAGAGAGCGAGACAGACATGAGTCTTCCACCCGCTCTCGAGGAGATCTTAC 218
QY 440 LysAspLeuTyrAlaThrAlaIleLeuThr 449
Db 217 AAGGACTTCTACTAGCAGAGAGTCTCTGAGC 188

RESULT 10
US-10-125-237-5/c
: Sequence 5, Application US/10125237
: Publication No. US20030022329A1
: GENERAL INFORMATION:
: APPLICANT: Tang, Y. Tom
: APPLICANT: Liu, Chenghua
```

```
: APPLICANT: Zhou, Ping
: APPLICANT: Asundi, Vinod
: APPLICANT: Ren, Feiyan
: APPLICANT: Zhao, Qing A.
: APPLICANT: Xue, Aidong J.
: APPLICANT: Xue, Jie
: APPLICANT: Zhang, Jie
: APPLICANT: Wehrman, Tom
: APPLICANT: Wang, Jian-Rui
: APPLICANT: Drmanac, Radote T.
: TITLE OF INVENTION: No. US20030022329A1e1 Nucleic Acids and
: FILE OF INVENTION: Polypeptides
: FILE REFERENCE: 791CIP2ADiv
: CURRENT APPLICATION NUMBER: US/10/125,237
: PRIOR FILING DATE: 2002-04-17
: PRIOR APPLICATION NUMBER: 09/668,317
: PRIOR FILING DATE: 2000-09-22
: PRIOR APPLICATION NUMBER: 09/552,929
: NUMBER OF SEQ ID NOS: 91
: SOFTWARE: pt_FL_genes Version 2.0
: SEQ ID NO 5
: LENGTH: 2307
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1612)..(2184)
US-10-125-237-5

Alignment Scores:
Pred. No.: 1,61e-232 Length: 2307
Score: 2065.50 Matches: 414
Percent Similarity: 85.71% Conservative: 6
Best Local Similarity: 84.49% Mismatches: 7
Query Match: 58.88% Indels: 63
DB: Gaps: 2

US-09-931-007A-1 (1-688) x US-10-125-237-5 (1-2307)
QY 23 AlaAsnIleSerGlnLumethrMetValAspThrGluMetProPheThrProThrAsn 42
Db 1657 GCCGCGGTGGCCGACGAATGACATGTTGACACAGAGATGCAATTCGGGCCACCAAC 1598
QY 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62
Db 1597 TTTGGATAGCTCCGTGATCTCTCCGTAATGGAAGACACTCCACCTCTTGTATATC 1538
QY 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82
Db 1537 AAGCCTTCACTAAGTGTGACTTCTCCACGATTTCTACTCCACATTCAGAACATTCGA 1478
QY 83 PheThrArgThrAspProValAlaAspTyrLysTyrAspLeuLysLeuGlnGluTyr 102
Db 1477 TTCACAGAGACAGATCCATGAGTGTGCGAGATTACAAAGTATGACCTGAACATTCAG 1418
QY 103 GlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeu 122
Db 1417 CAAAGGCAATCAAGTGTGAGGCTGCATCTCCACCTTATATTTCGAGAAAGCTCAGCTC 1358
QY 123 TyrAsn----- 124
Db 1357 TACAATTAAGCTCATGAGAGCCCTTCAACCTCTCATGCAATTTGAATGTGCTGTCTGT 1298
QY 124 ----- 124
Db 1297 GGAGATAAAGCTTCTGATTTCTATGAGATTCATGCTTGTGAAGATGCAAGGTTTC 1238
QY 124 ----- 124
Db 1237 TTCGCGAGACATCAGATTTGAGGCTTATCTATGACAGATGTCATTTACTGTGGATC 1178
QY 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140
|||||
```

Db 1177 CACAAAAAGTAGAATAAATGTCAGTACTGTGCGTTTCAGAAATGCTTGACAGTGGG 1118  
Qy MetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLysLeu 160  
Db 1117 ATGTCTCATATGCGATGAGGTTTGCGGAGTCCGACAGCGCCGAGAGAGAGAGTGTGTTG 1058  
Qy 161 AlaGluIleSerSerAspIleAspGlnLeuAsnProGluSerAlaAspLeuArg--Ala 179  
Db 1057 GCGGAGATCTCCAGATATGACAGCAGCTGAATCCAGATCCGCTGAGCTCCGTCAGCGCC 998  
Qy 180 LeuAlaIleHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAlaLysAla 199  
Db 997 CTGGGAAAAACATTGTATGACCATCATCAATAAAGTCTTCGCGTGACCAAGCAAGCGCG 938  
Qy 200 ArgAlaIleLeuThrGlyLysThrAspLysSerProPheValIleTyrAspMetAsn 219  
Db 937 AGGGGATCTTGACAGGAAGAAGCAACACAAATACCATTCGTTATGTATGACATGAT 878  
Qy 220 SerLeuMetMetGlyGluAspLysIleLysPheLysHisIleThrProLeuGlnGluGln 239  
Db 877 TCCTTAATGATGGAGAGATAAATCAAGTTCAAAACATCACCCCTCGAGAGAGCAG 818  
Qy 240 SerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGluAlaVal 259  
Db 817 AGCAAGAGAGGTGGCCATCGCATCTTCAGGGCTCCAGTTCCGCTCGTGAGGCTGTG 758  
Qy 260 GlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAsp 279  
Db 757 CAGGAGATCACAGAGTATGCCAAAAGCATTTCTGTTTGTAAATCTTGACTTGACAGAC 698  
Qy 280 GlnValIleThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeu 299  
Db 697 CAAGTAACCTCTCTCAAAATATGAGATCCACAGATCATTTACAAATGCTGGCCCTGTG 638  
Qy 300 MetAsnLysAspGlyValLeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeu 319  
Db 637 ATGAATAAAGATGGGGTCTCATATCCAGGGCCCAAGGCTTCATACAAAGGAGTTTCTA 578  
Qy 320 LysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAlaValLys 339  
Db 577 AAGAGCCCTGGCAAAACCTTTTGGTACTTTATGAGCCCAAGTTTGAATTTGCTGTAGAG 518  
Qy 340 PheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIleIleLeu 359  
Db 517 TTCATATGACATGGAATGTATGACAGCGACTTGGCAATATTATTCATTTCTC 458  
Qy 360 SerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeu 379  
Db 457 AGTGAGACCGCCCGAGTTTCTGTAATGTGAAGCCCATTTGAAGACATTCAGACAACTG 398  
Qy 380 LeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerSerGlnLeuPheAla 399  
Db 397 CTACAAAGCCCTGGAGCTCCAGCTGAGCTGAACACCCCTGAGTCTCCACAGCTGTTTCC 338  
Qy 400 LysLeuLeuGlnLysMetThrAspLeuArgGlnIleValIleThrGluHisValGlnLeuLeu 419  
Db 337 AAGCTGCTCCAGAAATATACAGACTCAGACAGATTTGCAGGAGACAGTGCAGCTACTG 278  
Qy 420 GlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGluIleTyr 439  
Db 277 CAGGATCAAGAAAGACGAGACACATGTGATTTCAACCGCTCCTCAGAGATGTTAC 218  
Qy 440 LysAspLeuTyrAlaIleThrAlaIleLeuThr 449  
Db 217 AAGGACTTGTACTAGCAGAGAGTCTCTGAGC 188

RESULT 11  
US-09-816-828-9/c

; Sequence 9, Application US/0916828

; Patent No. US20020150898A1

; GENERAL INFORMATION:

; APPLICANT: Tang, Y. Tom

; APPLICANT: Zhou, Ping

; APPLICANT: Goodrich, Ryle  
; APPLICANT: Asundi, Vinod  
; APPLICANT: Ren, Feiyan  
; APPLICANT: Xue, Aldong J.  
; APPLICANT: Ma, Yuning  
; APPLICANT: Wang, Zhimei  
; APPLICANT: Zhao, Qing A.  
; APPLICANT: Zhang, Jie  
; APPLICANT: Wang, Jian-Rui  
; APPLICANT: Drmanac, Radote T.  
; TITLE OF INVENTION: No. US20020150898A1el Nucleic Acids and  
; FILE REFERENCE: Polypeptides  
; CURRENT FILING DATE: 2001-03-22  
; PRIOR APPLICATION NUMBER: 09/770,160  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: 09/552,929  
; PRIOR FILING DATE: 2000-04-18  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: pl\_files Version 2.0  
; SEQ ID NO 9  
; LENGTH: 2329  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1612)..(2211)  
; US-09-816-828-9

#### Alignment Scores:

Pred. No.:	1.63e-232	Length:	2329
Score:	2065.50	Matches:	414
Percent Similarity:	85.71%	Conservative:	6
Best Local Similarity:	84.49%	Mismatches:	7
Query Match:	58.88%	Indels:	63
DB:	10	Gaps:	2

US-09-931-007a-1 (1-688) x US-09-816-828-9 (1-2329)

Qy 23 AlaAsnIleSerGlnGluMetThrMetValAspThrGluMetProPheThrProThrAsn 42  
Db 1657 GCCGCCGTGGCGGAGAAATGACCATGTTGACACAGAGATCGATTCGTGCCCCAAC 1598  
Qy 43 PheGlyIleSerSerValAspLeuSerValMetGluAspHisSerHisSerPheAspIle 62  
Db 1597 TTTGGGATCACCTCCGTGATCTCTCCGTAATGGAAGACACCTCCACTCTTGATATG 1538  
Qy 63 LysProPheThrThrValAspPheSerSerIleSerThrProHisTyrGluAspIlePro 82  
Db 1537 AAGCCCTTCACTACTAGTGTGACTTCTCCAGCATTTCTACTCCACATTAAGAAACATTCCA 1478  
Qy 83 PheThrArgThrAspProValValAlaAspTyrIleTyrIleAspLeuLysLeuGlnGluTyr 102  
Db 1477 TTCAAAAGAAAGATCCAGTGTTCAGATTACAGTATGACTGAACCTCAAGAGTAC 1418  
Qy 103 GlnSerAlaIleLysValGluProAlaSerProProTyrTyrSerGluLysThrGlnLeu 122  
Db 1417 CAAGTGAATCAAGATGAGAGCTGATCTCCACCTTAATTATTCGAGAAAGACTCAGCTC 1358  
Qy 123 TyrAsn----- 124  
Db 1357 TACATTAAGCCTCATGAAGAGCTTCAACTCCCATGCGCATTTGAATGTGCTGTCTGT 1298  
Qy 124 ----- 124  
Db 1297 GGAGATAAGCTTCTGATTTACATGAGACTGCATGCTGTGGAAGATGCAAGGGTTTC 1238  
Qy 124 ----- 124  
Db 1237 TTCCGGAGAACATTCAGTTGAGCTTATCTATGACAGATGATCTTAACCTCCGAGTC 1178  
Qy 125 -----ArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAlaValGly 140

```
Db 1177 CACAAAAAGTGAATAATCTAGTACTGCTCGCTTCAGAAAGCCCTGCAGTGGG 1118
Oy 141 MetSerHisasnAlaIleargPhegluArgMetProgluAlaIleuysgluylsleuLeu 160
Db 1117 ATCTCTCAATATGCCATCCAGTTTGGCGGATGCACAGGCCGACGAAGGGAAGCTGTTG 1058
Oy 161 AlaIleuIleSerSerAspIleasnProgluSerAlaIleuArg---Ala 179
Db 1057 GCGGACATCTCCAGTGATATCCAGCCAGTGAATCCAGAGTCCGCTGACCTCGTCAAGGCC 998
Oy 180 LeuAlaIleuHisleuTyrAspSerTyrIleLeuSerPheProleuThrIysAlaIysAla 199
Db 997 CTGGCAAAACATTGTTATGACTCATCAATAAAGTCCCTCCGCTGCACCAAGCAAGGGCG 938
Oy 200 ArgAlaIleleuThrgluylsThrThrAspIysSerProPheValIleTyrAspMetasn 219
Db 937 AGGGCATCTTGCACAGGAAGACACAGACAAATCACCATTCGTTATCTATGACATGAAT 878
Oy 220 SerLeuMetMetgluyluAspIysIleIysPheIysHisIleThrProleuIngIuIn 239
Db 877 TCCTTATGATGGAGAGATATAATCAAGTTCAAAACATCACCCCTCGCAGAGACAG 818
Oy 240 SerIysgluValAlaIleargIlePheIngIuIysgluPheargSerValIuAlaVal 259
Db 817 AGCAAGAGGTGGCCATCCGCACTTTCAGGGCTCCAGTTTCGCTCCGTCGAGCGCTGTG 758
Oy 260 GluIleIleThrIleTyrAlaIysSerIleProgluPheValAsnleuAspIleuAsnAsp 279
Db 757 CAGGACATCACAGATATGCCAAAGACATTCCTGGTTTGTAAATCTTGACTTGAAGCAG 698
Oy 280 GluValIleThrleuLeuIysTyrIleValHisgluIleIleTyrIleThrleuAlaSerleu 299
Db 697 CAAGTAACCTCTCTCAATATATGAGTCCAGAGATCATTTACACATTCGTCGCTCTTGG 638
Oy 300 MetAsnIlyAspIylValIleuIleSerIuIyIngIuIysPheMetThrargIuPheleu 319
Db 637 ATGAAATAAGATGGGTCTCATATCCGAGGGCCAGGCTTCATGCAAGGGAGTTTCTA 578
Oy 320 LysSerleuArgIysProPhegluAspPheMetgluProIysPhegluPheAlaValIys 339
Db 577 AAGAGCTCGCAAGGCTTTTGGTGCATTTATGAGCCCAAGTTTATGTTGCTGTGAAG 518
Oy 340 PheAsnAlaIleuIleuAspSerAspIleuAlaIlePheIleAlaValIleIleu 359
Db 517 TTCAATGCACTGGAATTAGATGACAGCACTTGGCAATATTATTCTCTGCATATTATTC 458
Oy 360 SerGlyAspArgProgluIleuAsnValIysProIleGlulAspIleGlulAspAsnleu 379
Db 457 AGTGGAGCCGCCCAAGTTTGGTGAATGTGAAGCCCATTTGAAGACATTCACAGCAACCTG 398
Oy 380 LeuGlnAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 399
Db 397 CTACAAGCCCTGGAGTCCAGCTGAAGCTGAACACCCCTGATCTCTCACACCTGTTTGGC 338
Oy 400 LysLeuLeuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 419
Db 337 AAGCTGCTCCAGAAATGACAGACCTCACAGATGTGTACAGGAACAGTGCAGCTACTG 278
Oy 420 GlnValIleIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 439
Db 277 CAGGTGATCAAGAGAGGAGAGACAGATGAGTCTTCACCCCTCTCTGACAGAGATCTAC 218
Oy 440 LysAspLeuTyrAlaIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleuIleu 449
Db 217 AAGGACTTGTACTAGCAGAGAGTCTCTGAGC 188
```

RESULT 12

US-10-142-373-1

; Sequence 1, Application US/70142373

; Patent No. US20020137665A1

; GENERAL INFORMATION:

; APPLICANT: EVANS, Ronald

```

; APPLICANT: FORMAN, Barry
; TITLE OF INVENTION: MODULATORS OF PEROXISOME PROLIFERATOR ACTIVATED
; TITLE OF INVENTION: RECEPTOR-GAMMA.
; TITLE OF INVENTION: AND METHODS FOR THE USE THEREOF
; FILE REFERENCE: SALK1480-2
; CURRENT APPLICATION NUMBER: US/10/142,373
; PRIOR APPLICATION NUMBER: US/09/788,070
; PRIOR FILING DATE: 2001-02-16
; PRIOR APPLICATION NUMBER: US 09/955,302
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 2005
; TYPE: DNA
; ORGANISM: Mus Musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (352)..(1776)
US-10-142-373-1

Alignment Scores:
Pred. No.: 1,48e-232 Length: 2005
Score: 2065.00 Matches: 408
Percent Similarity: 86.40% Conservative: 5
Best Local Similarity: 85.36% Mismatches: 3
Query Match: 58.87% Indels: 62
DB: Gaps: 1

US-09-931-007a-1 (1-688) x US-10-142-373-1 (1-2005)
Oy 28 GluMetThrMetValAspThrGluMetProPheThrProThrAsnPhegluIleSer 47
Db 343 GAAATTCACATGATGTTGACACACAGATGCCATTCCTGGCCACCAATTCGGAATGACTCT 402
Oy 48 ValAspLeuSerValMetGluAspHisSerHisSerPheAspIleuysProPheThrThr 67
Db 403 GTGACCTCTCCGTGATGAGAACACCTGCAATCTCTTGATCATCAAGCCCTTACACA 462
Oy 68 ValAspPheSerSerIleSerThrProHisTyrGluAspIleProPheThrargThrAsp 87
Db 463 GTTGATTTCTCCAGCATTTCTGCTCCACACTATGAAAGACATTCATCAAGACCTGAC 522
Oy 88 ProValAlaIleAspTyrIleTyrAspLeuIleuIleuIleuIleuIleuIleuIleuIleu 107
Db 523 CCAATGCTGTCGATTACAAATATGACCTGAAGCTCCAAAGATACCAAGTCCGATCAA 582
Oy 108 ValGluProAlaSerProProTyrTyrSerGluIysThrGlnLeuTyrAsn----- 124
Db 583 GTAGAACCCTGCATCTCCACTTATTATTCTGAAAGACCACCTCTACACAGGCTCAT 642
Oy 124 ----- 124
Db 643 GAAGAACCTTCTAACTCCCTCATGGCCATGAGTGCAGATCTGTGGGGAATAAGCATCA 702
Oy 124 ----- 124
Db 703 GGCTTCCACTATGAGATTCATGCTGTGAAGATCAAGGTTTTTTCGGAAGAACATC 762
Oy 125 ----- 125
Db 763 CGATTGAAGCTTATTATGATAGTGTGATCTTAATCGCCGATCCACAAAAAAGTGA 822
Oy 126 AsnIysCysGlnTyrCysArgPheGlnIlyCysLeuAlaValAlaGlyMetSerHisAsnAla 145
Db 823 AATTAATGTCAGTACTGTCGTTTCAGAGTGCCTTGGCTGGGATGTCTCACATCC 882
Oy 146 IleArgPheGlyIleArgMetProGlnAlaIleuIysgluIysIleuLeuAlaGluIleSerSer 165
Db 883 ATCAGGTTTGGGCGGATGCTCCAGAGGCCCAAGAGGAAAGAGTCTGTCGAGATCTCCAGT 942
Oy 166 AspIleasnProIleuAsnProIleuSerAlaIleuArgAlaIleuAlaIleuIleuIleuIleu 185
```

```
|||||
Db 943 GATATGACACGCTGACCCAGAGCTCTGATCTGCGACGCCCTTGCAAGACTTTGTAT 1002
Qy 186 ASPserTyrIleLysSerPheProLeuThrLysAlaLysAlaArgAlaIleLeuThrGly 205
Db 1003 GACTCATACATTAAGTCTCTCCGCTGACCAAGCCAAAGCGAGGCGCATCTTGACAGGA 1062
Qy 206 LysThrThrAspLysSerProPheValIleTyrAspMetLysSerLeuMetLeuGly 225
Db 1063 AAGACACGAGCAAAATCAACATTTTGTCTACATGATCAATGAATCTTATATGATGGAGAA 1122
Qy 226 AspLysIleLysPheLysHisIleThrProLeuGlnGlnGlnSerLysGlyValAlaIle 245
Db 1123 GATTAATAATCAAGTTCAACATATACCCCCCTGCGAGGACAGCAAGAGAGTGGCCATC 1182
Qy 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGlnIlePheThrLys 265
Db 1183 CGAATTTTCAAGGGTGGCGAGTTTGATCCGTAGAAAGCCGTGCAAGAGATCACAGAGTAT 1242
Qy 266 AlaLysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285
Db 1243 GCCAAAAATATCCCTGGTTTCATTAACTTGATTTGATGACCAAGTGACTCTGCTCAAG 1302
Qy 286 TyrGlyValHisGlnIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305
Db 1303 TATGCTGTCCATGAGATCATCTACACAGATGCTGGCCCTCGTATGAATAAGATGAGATC 1362
Qy 306 LeuIleSerGlnGlyGlnGlyPheMetThrArgGlnPheLeuLysSerLeuArgLysPro 325
Db 1363 CTCATCTCAGAGGGCCCAAGATTCATGACCAAGGAGTTCTCAAAAAGCCTGGGAAAGCCC 1422
Qy 326 PheGlyAspPheMetGlnProLysPheGlnPheAlaValLysPheAsnAlaLeuGlnLeu 345
Db 1423 TTGGTGACTTATAGACCCCTAAGTTTGAAGTTTCTGTAAGTTCAATGCACTGGAATTA 1482
Qy 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365
Db 1483 GATGACATGACTTGGCTATATTTTATCTGCTGATTAATTCACAGTGAGACCGCCACAGGC 1542
Qy 366 LeuLeuAsnValLysProIleGlnAspIleGlnAsnAsnLeuLeuGlnAlaLeuGlnLeu 385
Db 1543 TTGCTGAACGCTGAAGCCCATGACGAGACATCCAGACAAACCTGCTCAGGCCCCTGAACTG 1602
Qy 386 GlnLeuLysLeuAsnHisProGlnSerSerGlnLeuPheAlaLysLeuGlnLysMet 405
Db 1603 CAGCTCAAGCTGAATCACCCAGAGCTCTCTCAGCTTTCGCAAGAGTCTCCAGAAAGT 1662
Qy 406 ThrAspLeuArgGlnIleValIleThrLysIleValGlnLeuLeuGlnValIleLysLysThr 425
Db 1663 ACAGACCTCAGGACAGATGCTCACAGACAGCAGTGCAGCTACATGATCATCAAGAAAGACA 1722
Qy 426 GluThrAspMetSerLeuHisProLeuGlnGlnIleTyrLysAspLeuTyr 443
Db 1723 GAGACAGACATGAGCTTCCACCCCTGCTCCAGAGATCTACAAAGACTTTGTAT 1776

RESULT 13
US-09-954-456-554
: Sequence 554, Application US/09954456
: Patent No. US20020115057A1
: GENERAL INFORMATION:
: APPLICANT: Young, Paul
: TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
: TITLE OF INVENTION: Sets
: FILE REFERENCE: 689290-76
: CURRENT APPLICATION NUMBER: US/09/954,456
: CURRENT FILING DATE: 2001-09-18
: PRIOR APPLICATION NUMBER: US/60/233,617
: PRIOR FILING DATE: 2000-09-18
: PRIOR APPLICATION NUMBER: US/60/234,052
: PRIOR FILING DATE: 2000-09-20
: PRIOR APPLICATION NUMBER: US/60/234,923
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,134
```

```
: PRIOR FILING DATE: 2000-09-25
: PRIOR APPLICATION NUMBER: US/60/235,637
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,638
: PRIOR FILING DATE: 2000-09-26
: PRIOR APPLICATION NUMBER: US/60/235,711
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,720
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,840
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: US/60/235,863
: PRIOR FILING DATE: 2000-09-27
: NUMBER OF SEQ ID NOS: 2276
: SOFTWARE: PatentIn version 3.0
: SEQ ID NO 554
: LENGTH: 3301
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc-feature
: OTHER INFORMATION: n-a,t,g or c
US-09-954-456-554

Alignment Scores:
Pred. No.: 4,14e-119 Length: 3301
Score: 1101.50 Matches: 208
Percent Similarity: 80.98% Conservative: 56
Best Local Similarity: 63.80% Mismatches: 61
Query Match: 31.40% Indels: 1
DB: caps: 1

US-09-931-007a-1 (1-688) x US-09-954-456-554 (1-3301)
Qy 119 LysThrGlnLeuTyrAsnArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
Db 683 AAGATTCAAGAAAGAAAGAACCGGCAAGAGTGCAGTACTCCGCTTCCAAAGATGCTTGCA 742
Qy 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGlnLysGlnLys 158
Db 743 CTGGGCATGTACACACAGCCTATCCGTTTGGTCGATGCGGAGGCTGAGAGAGAGAG 802
Qy 159 LeuLeuAlaGlnIle---SerSerAspLysAspGlnLeuAsnProGlnLysAspLeu 177
Db 803 CTGCTGGCAGAGGCTGTACGCAAGACGAGGAGGCCAGTACAAACCCACAGGTGGCGACCTG 862
Qy 178 ArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
Db 863 AAGGCTTCTCCAAACACATCTACATGCTTACCTGAAACCTTCACATGACCAAAAAG 922
Qy 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
Db 923 AAGGCCCGCAGATCTCTACCGGCAAGACCAACGACAGGCGCCCTTGTATCCACAGAC 982
Qy 218 MetAsnSerLeuMetMetGlyLysAspLysIleLysPheLysHisIleThrProLeuGln 237
Db 983 ATCGACAGATGTGGCAGCAGAGAGAGAGGCGTGGTGGAAGACAGATGGTAATGGCCTG 1042
Qy 238 GlnGlnSerLysGlnValAlaIleArgIlePheGlnGlnGlyCysGlnPheArgSerValGlu 257
Db 1043 CCTCCCTCAAGAGATAGGCTGCAGCTCTTACCCCTCCAGTGCACCAAGTGGAG 1102
Qy 258 AlaValGlnGlnIleThrGlnTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277
Db 1103 ACCGTCGGGAGCTCACAGATGTCGCCAAGAGCATCCCACTCTACAGAGCTCTTCTC 1162
Qy 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGlnIleIleTyrThrMetLeuAla 297
Db 1163 AAGCAGCAGGTACCTCTCTCAAGATAGCGTGACAGCAGGAGCATCTTGGCATGCTGCGC 1222
Qy 298 SerLeuMetAsnLysAspGlyValLeuLeuSerGlnGlnGlnGlyGlnGlyPheMetThrArgGlu 317
Db 1223 TCTATCTCAACAAAGAGAGGAGGCTGCTGTGATGACCAACGACGACGTGCTTGTACCCGCTGAG 1282
```



```
OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
      |||||.....|||
Db 1283 TTCCTGCGAGCGCTCGCAACCCCTTCAGTATCATTTGAGCGCTTAAGTTGATTTGCT 1342

OY 338 ValLysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIle 357
      |||||.....|||
Db 1343 GTCAACCTTCACGCCCTCGAAGCTTGATGACAGGACCTGGCCCTATTCATTGCGGCCATC 1402

OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluSplIleGlnAsp 377
      |||||.....|||
Db 1403 ATTCTGTGGAGACCGCGCCAGGCTCATGAACTTCACCGGGTGAAGCTATCCAGGAC 1462

OY 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerGlnLeu 397
      |||||.....|||
Db 1463 ACCATCTCGTGGAGACCGCGCCAGGCTCATGAACTTCACCGGGTGAAGCTATCCAGGAC 1522

OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
      |||||.....|||
Db 1523 TTCGCCAACCTGCTGCAAGATGGCTGACCTGGCGCACTGCTCACCGAGACGCCCCAG 1582

OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437
      |||||.....|||
Db 1583 AGGATGCAAGCGGATCAAGAGACCCGAACCGAGACTCGCTGCACCCCTGCTCCAGGAG 1642

OY 438 IleTyrLysAspLeuTyr 443
      |||||.....|||
Db 1643 ATCTACAGACATGTATC 1660

RESULT 14
US-10-013-807-1
; Sequence 1, Application US/10013807
; Patent No. US20020112256A1
; GENERAL INFORMATION:
; APPLICANT: Genether, Catherine
; APPLICANT: Phillips, Russell
; APPLICANT: Allen, Keith D.
; TITLE OF INVENTION: TRANSGENIC MICE CONTAINING
; TITLE OF INVENTION: PROLIFERATOR-ACTIVATED RECEPTOR GENE DISRUPTIONS
; FILE REFERENCE: R-683
; CURRENT APPLICATION NUMBER: US/10/013,807
; CURRENT FILING DATE: 2001-12-11
; PRIOR APPLICATION NUMBER: US 60/254,916
; PRIOR FILING DATE: 2000-12-11
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-013-807-1

Alignment Scores:
Pred. No.: 8.76e-117 Length: 1323
Score: 1076.50 Matches: 203
Percent Similarity: 80.06% Conservative: 58
Best Local Similarity: 62.27% Mismatches: 64
Query Match: 30.69% Indels: 1
DB: 12 Gaps: 1

US-09-931-007a-1 (1-688) x US-10-013-807-1 (1-1323)

OY 119 LysThrGlnLeuTyrAsnArgAsnLysCysGlnTyrCysArgPheGlnLysCysLeuAla 138
      |||||.....|||
Db 343 AAGATCCCAAGAAAGACCGCAAGATGTCTGACTGCGCTTCAGAGAGTCCCTGGCA 402

OY 139 ValGlyMetSerHisAsnAlaIleArgPheGlyArgMetProGlnAlaGluLysGluLys 158
      |||||.....|||
Db 403 CTGCGCATGTCCGACAAACCGCTTCCGCTTGGACGATGCCGAGCGCAAGAGAGAG 462

OY 159 LeuLeuAlaGluIle---SerSerAspIleAspGlnLeuAsnProGluSerLysLeu 177
      |||||.....|||
Db 463 CTGGTGGGGGCTGACTGCTCAGGAGGGGGTCCAGACACACCCCACTGGCTGAGCTG 522
```

```
OY 178 ArgAlaLeuAlaLysHisLeuTyrAspSerTyrIleLysSerPheProLeuThrLysAla 197
      |||||.....|||
Db 523 AAGCGCTTCTTAAGACATCTACAAAGCGCTACCTGAAATACTTACATGACCAAAAG 582

OY 198 LysAlaArgAlaIleLeuThrGlyLysThrThrAspLysSerProPheValIleTyrAsp 217
      |||||.....|||
Db 583 AAGCGCGGAGGATCTCTCCCGCAAGTCCACCAACGACCCCTTGTCTTCCAGCAG 642

OY 218 MetAsnSerLeuMetMetLysGluAspLysIleLysPheLysHisIleThrProLeuGln 237
      |||||.....|||
Db 643 ATCGAGACATGTGGCAGCAAGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 702

OY 238 GluGlnSerLysGluValAlaIleArgIlePheGlnGlyCysGlnPheArgSerValGlu 257
      |||||.....|||
Db 703 CGGCCCTACACGAGATCATGTTGTCACAGTCTTACCGCTGCCAGTCCACCACTGGAG 762

OY 258 AlaValGlnGluIleThrGluTyrAlaLysSerIleProGlyPheValAsnLeuAspLeu 277
      |||||.....|||
Db 763 ACAGTCCGAGACTCACCGAGTTCCCAAGAACATCCCACTTCCAGAGCTCTTCTC 822

OY 278 AsnAspGlnValThrLeuLeuLysTyrGlyValHisGluIleIleTyrThrMetLeuAla 297
      |||||.....|||
Db 823 AATGACAGGTACCCCTCTCAAGTATGGCTGCAAGAGCCATCTTGCCATGCTGGCC 882

OY 298 SerLeuMetAsnLysAspGlyValLeuLysSerGluGlnGlyGlnGlyPheMetThrArgLys 317
      |||||.....|||
Db 883 TCCATGCTCAACAAAGAGCGGCTGCTGTGCGCCACAGGCAAGTGGCTTGTCACCCAGAG 942

OY 318 PheLeuLysSerLeuArgLysProPheGlyAspPheMetGluProLysPheGluPheAla 337
      |||||.....|||
Db 943 TTCCTGCGAGCTCGCGAAGCCCTTCAGTGCATCATTTGAGCCCAAGTGGAGTTGCT 1002

OY 338 ValLysPheAsnAlaLeuGluLeuAspSerAspLeuAlaIlePheIleAlaValIle 357
      |||||.....|||
Db 1003 GTCAAGTTCATAGCCCTGAGCTGCATGACACTGAGCTTGCGCTTTCATGCGGCCATC 1062

OY 358 IleLeuSerGlyAspArgProGlyLeuLeuAsnValLysProIleGluAspIleGlnAsp 377
      |||||.....|||
Db 1063 ATTCTGTGGAGACCGCGCCAGGCTCATGAAATGTGCCAGAGTGAAGCATCCAGAGAC 1122

OY 378 AsnLeuLeuGlnAlaLeuGluLeuGlnLeuLysLeuAsnHisProGluSerGlnLeu 397
      |||||.....|||
Db 1123 ACCATCTCGCGGCTCGAAGATTCATCATCTGCGAGGACCAACCCGAGACGAGTACCTC 1182

OY 398 PheAlaLysLeuLeuGlnLysMetThrAspLeuArgGlnIleValThrGluHisValGln 417
      |||||.....|||
Db 1183 TTCGCCAACCTGCTCAGAGATGGCAGACCTCGCGGACGTGTCTGAGACATGCCAG 1242

OY 418 LeuLeuGlnValIleLysLysThrGluThrAspMetSerLeuHisProLeuLeuGlnGlu 437
      |||||.....|||
Db 1243 ATGATGCAAGTGTCAAGAGAGCGAGAGTGAAGACTGCTGTGCACCCCTGCTCCAGGAA 1302

OY 438 IleTyrLysAspLeuTyr 443
      |||||.....|||
Db 1303 ATCTACAGACATGTATC 1320

RESULT 15
US-10-033-528-1855
; Sequence 1855, Application US/10033528
; Patent No. US20020131971A1
; GENERAL INFORMATION:
; APPLICANT: King, Gordon E.
; APPLICANT: Meagher, Madeleine Joy
; APPLICANT: Xu, Jianshun
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF COLON CANCER
; FILE REFERENCE: 210121.547C1
; CURRENT APPLICATION NUMBER: US/10/033,528
; CURRENT FILING DATE: 2001-12-26
; NUMBER OF SEQ ID NOS: 1896
; SOFTWARE: FastSeq for Windows Version 4.0
```



; SEQ ID NO 1855  
; LENGTH: 428  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 5  
; OTHER INFORMATION: n = A,T,C or G  
US-10-033-528-1855

Alignment Scores:  
Pred. No.: 1.43e-73 Length: 428  
Score: 702.00 Matches: 140  
Percent Similarity: 99.29% Conservative: 0  
Best Local Similarity: 99.29% Mismatches: 1  
Query Match: 20.01% Indels: 0  
DB: 12 Gaps: 0

US-09-931-007a-1 (1-688) x US-10-033-528-1855 (1-428)

OY 246 ArgIlePheGlnGlyCysGlnPheArgSerValGluAlaValGlnGluIlePheGlnTyr 265  
|||  
DB 6 CCATCTTTCAGGCTGCCAGTTGCTCGTGAGGCTGTGACAGATCAGAGATAT 65  
OY 266 AlaIysSerIleProGlyPheValAsnLeuAspLeuAsnAspGlnValThrLeuLeuLys 285  
|||  
DB 66 GCCAAAGCATTCCTGCTTGTAAATCTGACTGAACGACCAAGTAACCTCTCTCAA 125  
OY 286 TyrGlyValHisGluIleIleTyrThrMetLeuAlaSerLeuMetAsnLysAspGlyVal 305  
|||  
DB 126 TATGGAGTCCAGAGATCATTTACCAATGCTGGCTCTTGATGATGAATAAAGATGGGCTT 185  
OY 306 LeuIleSerGluGlyGlnGlyPheMetThrArgGluPheLeuLysSerLeuArgLysPro 325  
|||  
DB 186 CTCATATCCAGAGGCCAGAGCTTTCATGACAAAGGAGTTTCTAAAGAGCTCGCAAGGCTT 245  
OY 326 PheGlyAspPheMetGluProLysPheGluPheAlaValLysPheAsnAlaLeuGluLeu 345  
|||  
DB 246 TTTGGTGACTTATGAGCCCAAGTTGAGTTGCTGTGAAGTTCATGCACTGGAATTA 305  
OY 346 AspAspSerAspLeuAlaIlePheIleAlaValIleIleLeuSerGlyAspArgProGly 365  
|||  
DB 306 GATGACAGCGACTTGCAATATTATTCCTGTCATATTCTCAGTGAGACCGCCAGGT 365  
OY 366 LeuLeuAsnValLysProIleGluAspIleGlnAspAsnLeuLeuGluAlaLeuGluLeu 385  
|||  
DB 366 TTGCTGAATGTAAGCCCATTTGAAGACATTCAGACAACTGCCACAAGCCCTGGAGCTC 425  
OY 386 Gln 386  
|||  
DB 426 CAG 428

Search completed: February 25, 2003, 06:18:30  
Job time : 293 secs

**THIS PAGE BLANK (USPTO)**